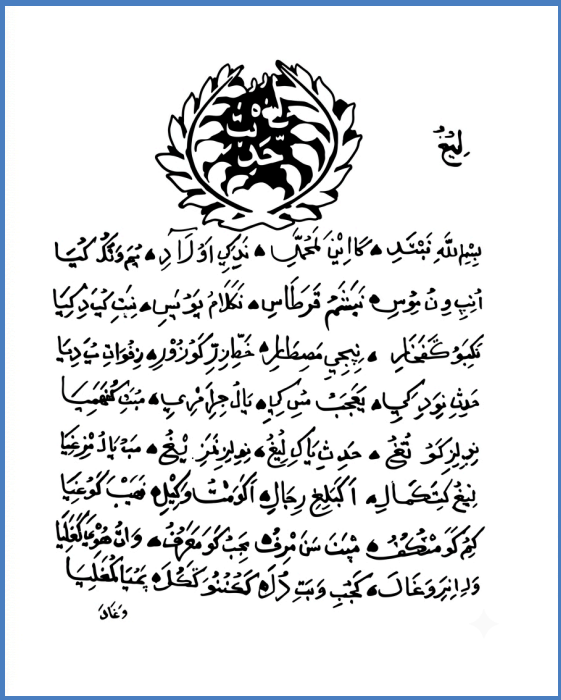
THE PROCRUSTEAN BED

The Fallacy of Accentual Metrics in Swahili Poetry

A Manifesto for Data-Driven Models for the Utenzi, Wimbo, and Shairi Meters



MATHIEU ROY

The Procrustean Bed: The Fallacy of Accentual Metrics in Swahili Poetry

A Manifesto for Data-Driven Models for the Utenzi, Wimbo, and Shairi Meters

Mathieu Roy, PhD

DL$_2$A - Buluu Publishing

DL_2A - Buluu Publishing

66, avenue des Champs-Élysées,

Paris, France.

www.dl2a.eu | contact@dl2a.eu

Distribution Model & Open Science Commitment: This publication operates on a dual-distribution model to ensure maximum accessibility for the global research community while sustaining independent publishing.

Open Access Digital Edition:

The intellectual content (text and figures) of this work is freely available as an author manuscript (PDF) and is distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0). You are free to read, download, copy, distribute, and adapt this material in any medium or format, even for commercial purposes, provided you give appropriate credit to the original author, provide a link to the license, and indicate if changes were made.

Digital Object Identifier: 10.5281/zenodo.21720211

Commercial Printed Edition:

Copyright © 2026 Mathieu Roy | ORCID 0000-0001-7684-4355

While the underlying text and figures are distributed under an Open Access license, the specific typographic arrangement, formatting, and cover design of the printed commercial edition remain the exclusive property of the publisher. Reproduction of the physical printed book or its specific layout without prior written permission is prohibited.

ISBN 978-1-291-70670-3

Library of Congress Control Number: 2026919513

Library of Congress Subject Headings: Swahili language--Versification | Swahili poetry--History and criticism. | Poetics. | Metrics and rhythmics.

Neno moja tukumbuke katika utungaji, nalo ni hili ya kuwa mawazo ya mtungaji yazue vina, sio vina kuzua mawazo.
— Sheikh Kaluta Amri Abedi, 1954, p. 1
Let us remember one thing in composition, and it is this: the thoughts of the composer must generate the rhymes, not the rhymes generating the thoughts.

Hapana shaka kuwa mtungaji mwenye busara ataona kuwa upo uhuru mwingi katika kutunga kwa vina na mizani.
— Ibrahim Noor Shariff, 1988, p. 48
There is no doubt that a wise composer will see that there is a great deal of freedom in composing with rhymes and syllabic measures.

Abstract

For more than a century, one lineage in the study of classical Swahili poetry has rested on an epistemological fallacy: forcing an autonomous poetic tradition onto the Procrustean bed of exogenous prosodic frameworks. This school holds that the line is generated by accentual beats or metrical feet: Edward Steere's anapaests (1870), Carl G. Büttner's iambs and trochees (1894), Alice Werner's two beats per line (1917–20), Clarissa Vierke's *Grenzschema*, an ideal two-stress template (2011), and Emiliano Minerba's trochaic innovation (2025). This book names the assumption behind that lineage the Exo-African Postulate: that a demanding African poetic architecture must owe its generative principle to a foreign model. Its cost has been to look past the explicit accounts of the Swahili-language poets and the material reality of the texts.

To test this lineage, this study formalizes the syllabic system the master poets describe — *vina* (the rhyme) and *mizani* (the syllable count) — as matrices and algebraic equations for the *utenzi*, *wimbo*, and *shairi* meters, and reduces each accentual claim to conditions precise enough to fail. A 3-Layer Theory, grounded in graphic reason, separates the syllabic law, the architectures built upon it, and their sung performance, placing lexical stress outside all three: the accentual school took a property of the Swahili language for a property of the meter. The poets' own metrical vocabulary points the same way: it borrowed technical terminology from Arabic, but never a word for a foot or a beat.

Both models then face the same eight conditions of falsifiability, four for each, across eight corpora of classical and contemporary poetry. Every line is read by rule from the text and by ear from a formal reading, with instrumental measurement applied to a hostile sample. Two control corpora of ordinary prose provide the baseline.

The syllabic model survives every test. The accentual model does not: across hundreds of tested lines, its predicted metrical feet fail to operate as a generative rule, appearing only where word choice happens to produce them. Its ideal configuration appears in ordinary prose whenever the words happen to supply it, and a poem written to satisfy it succeeds only by restricting its vocabulary to four-syllable words. Classical Swahili meter is governed by *vina* and *mizani*. This book does not discover that system; it falsifies the model imported to replace it. Every recording, script, and dataset is deposited in open access.

ABSTRACT **5**

INTRODUCTION: THE ROOT MISCONCEPTION OF EXOGENOUS ACCENTUAL METRICS **23**

QUICK REFERENCE GUIDE: MATRICES AND ALGEBRAIC EXPRESSIONS OF SWAHILI CLASSICAL METERS **33**

QRG.1. The Symbolic Alphabet 33
QRG.2. Utenzi Meter 36
QRG.2.1. Standard Matrix and Algebraic Equation for One Stanza (ubeti) *36*
QRG.3. Wimbo Meter 37
QRG.3.1. The Standard Bicolumnar Matrix and Algebraic Equation (Symmetrical and Asymmetrical) *37*
QRG.3.2. The Rupture Matrix: Vertical Asymmetry *38*
QRG.3.3. The Monocolumnar Matrix (Wimbo wa Zina Zitoto) *39*
QRG.3.4. The Hyper-Fractionated Matrices *40*
QRG.4. Shairi Meter 41
QRG.4.1. 16-Syllable Shairi Meter: Generative Matrix for One Stanza (ubeti) *41*
QRG.4.2. Shairi Meter: General Generative Matrix for One Stanza (ubeti) *42*
QRG.4.3. Shairi Meter: The Stanzaic Chaining Matrix (Mashairi ya Vidato) *43*

CHAPTER 1. EPISTEMOLOGICAL PRELIMINARIES: DATASETS AND THE EXO-AFRICAN POSTULATE **45**

1.1. The Chronological Division and the Scope of the Critique 45
1.2. The Methodological Imperative of Formal Modeling and Experimental Test 49
1.3. The Dammann Corpus: Historical Context, Ideological Bias, and the Neutrality of Raw Data 51
1.4. The Great Zimbabwe Syndrome and the Exo-African Postulate 54
1.5. From Early Heuristics to the Analytic Turn in Swahili Metrics 55

CHAPTER 2. THE ACCENTUAL SCHOOL OF CLASSICAL SWAHILI POETRY **59**

2.1. Phonological Grounds for Accentual Metrics in Swahili 59
2.2. The Accentual School: A Century of Exogenous Projection 62
2.2.1. The Early Pioneers (1870–1920) *62*
2.2.2. The Continuators: Consolidating the Accentual Tradition (1970s) *64*

2.2.3. Metrical Excursus: The Anatomy of Exogenous Feet *65*
2.3. CLARISSA VIERKE AND THE APPLICATION OF THE *GRENZSCHEMA* 69
2.3.1. Visualizing the Accentual Edifice: Key Concepts of Clarissa Vierke's Metrical Theory (2011) *70*
1. Key Metrical Concepts 72
1.a. Bipartite Two-Beat Accentual Model (Based on Alice Werner) . 72
1.b. Trochaic Nucleus Structure 72
1.c. Pre-Nuclei 72
1.d. The *Grenzschema* 72
1.e. Mixed Metrics (*gemischt-quantifizierend*) 72
2.3.2. The Mechanics of the "Swahili Foot" and the Trochaic Hypothesis *72*
2.3.3. Theoretical Adjustments: Arabic Cores and "Irrelevant" Syllables *75*
2.3.4. The Grenzschema: Evaluating the Abstract Cognitive Prototype *77*
2.3.5. The Methodological Shift Toward Performance and the Musical Rationale *82*
2.4. EMILIANO MINERBA AND THE COMPARATIVE APPROACH: THE DIFFUSIONIST HYPOTHESIS 84
2.4.1. A Necessary Preliminary: Reading Minerba as He Actually Is. *85*
2.4.2. The Limits of Extrapolation: Wolofal, Classical Swahili Poetry, and the Risk of a Double Erasure under Arabic Projection *87*
2.4.3. The Ḥumaynī Thesis: Surface Isomorphism vs. Deep Metrical Matrix *89*
1. Mapping the Classical Triad: The Proposed Yemeni Ancestors of Utenzi, Wimbo, and Shairi 89
2. Beyond Surface Isomorphism: Structural Counter-Evidence in Swahili Poetic Architecture 91
2.4.4. The "Trochaic Innovation": Structural Inevitability vs. Generative Intent *93*
1. Drawing the Metrical Target: The "Creolized Trochee" as a Texas Sharpshooter Fallacy 95
2. Disentangling Endogenous Mechanics, Macroscopic Interfaces, and Performance: Introducing the 3-Layer Framework 98
2.5. FORMALIZING THE ACCENTUAL MODELS FOR EXPERIMENTAL TESTING .. 102
2.5.1. Clarissa Vierke: The Syllabic Baseline and the Accentual Grenzschema *102*
1. Stratum 1: The Syllabic Baseline 103
2. Stratum 2: The Accentual Superstructure 103
2.a. The Symbolic Alphabet 104
2.b. Alice Werner's Bipartite Legacy: The Axioms Defining Two Accentual Groups Per Line 104
2.c. The Ideal Matrix (The *Grenzschema*) 105

3. Epistemological Parameters: The Conditions of Falsifiability 106
3.a. The Violation of the Two-Beat Axiom ($\boldsymbol{G \neq 2}$) 106
3.b. Testing the *Grenzschema* as a Statistical Tendency 106
3.c. The Arbitrariness of Asymmetrical Distributions 107
2.5.2. Formalizing the Minerba Accentual Model for Experimental Testing 107
1. The Minerba Matrix (The Creolized Trochaic Grenzschema) 108
1.a. The Octosyllabic Utenzi Line ($\boldsymbol{S = 8}$) and Shairi Hemistich ($\boldsymbol{H1 = H2 = 8}$) Matrix 109
1.b. The Dodecasyllabic Wimbo Matrix: The $\boldsymbol{4 + 8}$ Division ($\boldsymbol{S = 12}$) 112
2. Epistemological Parameters: Falsifying the "Trochaic Innovation" 113

CHAPTER 3. GRAPHIC REASON, SWAHILI PHONOLOGY, AND THE 3-LAYER THEORY OF SWAHILI POETICS 115

3.1. "*Wimbo Hutoka Ngomani*": From the Head to the Graphic Grid ... 115
3.2. Graphic Reason and the Objectification of the Verse 120
3.3. The 3-Layer Theory of Swahili Poetics 122
3.3.1. Layer 1: The Formal Foundation (The Syllabic Law) 122
3.3.2. Layer 2: The Architectural Interface and the Paradox of Hyper-Sophistication 124
3.3.3. Layer 3: Melodies, Mahadhi, and the Performative Fallacy ... 127
1. Alice Werner's Error (1917-20) and the Accentual Legacy 128
2. The One-to-Many Mapping: One Line, One Lexical Accentuation, Many Performances 132
3. The Written Contract: Bridging the Poet (mtunzi) and the Performer (mghani) in Layer 3 134
4. The Two Seats of the Formal Layers 135
3.3.4. Locating the Scholarship 137
3.4. Conclusion 140

CHAPTER 4. THE TERMINOLOGICAL PROOF: THE ACCENTUAL LEXICAL VOID AMID THE SWAHILI METRICAL LEXICON 141

4.1. The Architecture of Compositions: An Endogenous Taxonomy ... 141
4.1.1. The Arithmetic of the Measure: The Counting Paradigm 142
Mizani (Class 9/10) (The Measure) 142
Hisabu (Class 9/10) (The Count) 142
Urari (Class 11) (The Balance) 142
Silabi (Class 9/10) (The Syllable) vs. Sauti (Class 9/10) (The Voice). 143
Ta[*] (Class 9/10) / U[*] (Class 11) (The Stanzaic Line Counters) 144
4.1.2. The Structural and Cognitive Architecture 145
Aina (Class 9, pl. aina Class 10) (The Type / The Topological Variation) 145
Bahari (Class 9, pl. bahari Class 10) (The Meter / The Sea) 145

Definition 1: The Overarching Meter 145
Definition 2: The Omni-Stanzaic Rhyme and the "Habari" Shift 146
Kifungu (Class 7, pl. vifungu Class 8) (The Section) 147
Kipande (Class 7, pl. vipande Class 8) (The Segment / The Hemistich) 147
Mkondo (Class 3, pl. mikondo Class 4) (The Sub-Meter / The Syllabic Current) 148
Mshororo (Class 3, pl. mishororo Class 4) vs. Mstari (Class 3, pl. mistari Class 4) — The Structural Line vs. The Graphic Line 149
Ubeti (Class 11, pl. beti Class 10) (The Stanza) 150
Tungo (Class 9, pl. tungo Class 10) (The Composition) and Diwani (Class 9, pl. diwani Class 10) (The Anthology) 150
4.2. The Metric Locks, Boundaries, and Trajectories 151
Kina (Class 7, pl. vina Class 8) (The Rhyme) 152
Kiwango (Class 7, pl. viwango Class 8) (The Quantity / The Limit) 153
Kituo (Class 7, pl. vituo Class 8) vs. Kibwagizo (Class 7, pl. vibwagizo Class 8) (The Stop vs. The Refrain) 154
Definition 1: Kituo as the Metrical Pause or Caesura 154
Definition 2: Kituo vs. Kibwagizo (The Terminal Line vs. The Refrain) 154
4.2.1. The Topological Motifs of Rhyme Chains 155
Kupindulia (Class 15) (The Permutation / The Inversion) 156
Mshuko (Class 3, pl. mishuko Class 4) (The Descent) 156
Kurejelea (Class 15) (The Recurrence) 156
Mtiririko (Class 3, pl. mitiririko Class 4) (The Continuous Flow) 157
Kidato (Class 7, pl. vidato Class 8) (The Step / The Notch) 158
4.3. The Paradox of Arudhi and Ubeti: Arabic Shells for a Swahili Metrical Core 159
4.3.1. Ubeti: The Ultimate Lexical Shell 160
4.3.2. The Architectural Metamorphosis 160
4.3.3. The Illusion of a Total Borrowing 161
4.4. The Epistemological Arbiter: *Guni* (The Defective Verse) 163
4.5. The Division of Labor: Composition vs. Performance 164
4.5.1. The Architects of the Compositions 164
Mtunzi (Class 1, pl. watunzi Class 2) (The Composer / The Poet-Architect) 164
Mshairi (Class 1, pl. washairi Class 2) (The Poet) 165
Malenga (Class 1, pl. malenga Class 2) (The Master Poet / The Virtuoso) 166
4.5.2. The Performative Realm (The Execution) 166
Mghani (Class 1, pl. waghani Class 2) (The Performer / The Singer) . 166
Mahadhi (Class 6) (The Melody / The Performative Rhythm) 166
4.6. The Endogenous Accentual Lexical Void and Academic Projections 167

CHAPTER 5. THE EMPIRICAL REALITY OF *VINA NA MIZANI*: VOICES OF THE MASTER POETS OF THE SWAHILI LANGUAGE..... 171

5.1. Sheikh Kaluta Amri Abedi and the Absolute Syllabic Foundation (1954) 172
5.1.1. The Technical Definition of the Utenzi Parameters 173
5.1.2. The Ostensive Definition: Regulated Metrical Variations 174
5.1.3. The Syllable as the Sole Metric Unit 175
5.1.4. An Autonomous Intellectual Weapon 176
5.2. Shaaban Robert and the Ontological Necessity of the Syllabic Matrix 177
5.2.1. Mashairi ya Mizani: The Absolute Measure 178
5.2.2. The Architecture of Vina and the Habari of the Utenzi 179
5.2.3. The Epistemological Boundary: Guni as a Formal Defect 180
5.3. Ibrahim Noor Shariff: The Defense of Endogenous Sovereignty (1988) 181
5.3.1. The Lexical Architecture: Vina, Mizani, and Kiwango as an Indivisible Concept 185
5.3.2. Shariff's Definition of the Classical Meters: Utenzi, Wimbo, and Shairi 188
5.4. Ahmed Sheikh Nabhany and the Exhaustive Taxonomy of Classical Forms (2006) 194
5.4.1. The Three Primary Models: Shairi, Wimbo, and Utendi 196
1. Shairi 196
2. Wimbo 197
3. Utendi (Utenzi) 197
5.5. Hashil Seif Hashil: Acknowledging the Boundaries and the Formal Rupture (2009) 199
5.6. Mathias E. Mnyampala: The Generative Expansion of the Formal Syllabic Grid (1963–1971) 202
5.6.1. The Mtiririko (The Flow) 204
5.6.2. The Msisitizo (The Emphasis) 205
5.6.3. The Ngonjera (The Disputatio) 206
Focus 1: Charles M. Mnyampala and the Preservation of the Archive 207
5.7. Conclusion: Listening to Swahili Voices 208

CHAPTER 6. THE MARGINALIZED LEGACY: HOW THE ACCENTUAL SCHOOL OVERLOOKED THE SYLLABIC PRINCIPLES OF SWAHILI METRICS IN JAN KNAPPERT AND LYNDON HARRIES 211

6.1. The Francophone Epistemological Interlude: Charles Sacleux, Kris Berwouts and Contemporary Empirical Research 211

6.2. JAN KNAPPERT AND THE PRESERVATION OF THE *VINA NA MIZANI* 219
6.2.1. The Empirical Evidence: The Absolute Syllabism in Knappert's Meta-Vocabulary *223*
6.3. LYNDON HARRIES AND THE SYLLABIC ARCHITECTURE OF CLASSICAL SWAHILI POETRY 228
6.3.1. Lyndon Harries and the Recognition of Sui Generis Syllabism *230*
6.4. THE MECHANICS OF A SCHOLARLY MARGINALIZATION OF VINA NA *MIZANI* 233

CHAPTER 7. THE COST OF FORMAL AND SCIENTIFIC METHODS: EPISTEMOLOGICAL RUPTURE AND ACADEMIC INERTIA 241

A NOTE ON INSTITUTIONAL RECEPTION 249

CHAPTER 8. THE SOVEREIGN GRID: THE ENDOGENOUS LOGIC OF SWAHILI SYLLABIC METER 251

8.1. THE EPISTEMOLOGICAL FRAMEWORK: FROM DATA TO FORMALIZATION AND TESTS 251
8.2. FORMAL NOTATION AND SYMBOLIC ALPHABET OF THE METRICAL GRID. 252
8.3. THE UTENZI METER 256
Focus 2: Utenzi and the Poetics of "Doing" — Translating the Word Utenzi in English *257*
8.3.1. Utenzi Meter: Generative Matrix for One Stanza (ubeti) *259*
8.3.2. The Evolution of Graphic Reason: Two Spatial Instantiations *259*
1. The Ajami Manuscript Matrix 260
2. The Romanized Print Matrix 261
2.a. Style A (*mtindo* A) — *Utenzi* with two lines *(mishororo)* 262
2.b. Style B (*mtindo* B) — *Utenzi* with four lines *(mishororo)* 262
8.3.3. Methodological Note on the Data-Driven Model *263*
8.4. THE *WIMBO* METER 264
8.4.1. The Standard Bicolumnar Matrix and Algebraic Equation (Symmetrical and Asymmetrical) *265*
8.4.2. The Rupture Matrix: Vertical Asymmetry *265*
8.4.3. The Monocolumnar Matrix (Wimbo wa Zina Zitoto) *266*
8.4.4. The Hyper-Fractionated Matrices *266*
8.5. THE SHAIRI METER: FROM THE 16-SYLLABLE STANDARD TO THE GENERAL FORMAL GRID 267
8.5.1. 16-Syllable Shairi Meter: Generative Matrix and Algebraic Equation for One Stanza (ubeti) *268*
8.5.2. Shairi Meter: General Generative Matrix for One Stanza (ubeti) *269*

8.5.3. The Evolution of Graphic Reason: Two Spatial Instantiations 270
1. The Ajami Manuscript Matrix 270
2. The Romanized Print Matrix 271
2.a. The Canonical 16-Syllable Shairi 271
2.a.a. Kupindua kina kimoja — permutation of a single rhyme (b only) 272
2.a.b. Kupindua — permutation (rhyme a and b) 272
2.a.c. Kushuka — Descending: the rhyme chains a and b continue without permutation in the final line 273
2.a.d. Kurejelea — Recurrence: the final line repeats throughout the poem 273
2.b. Other traditional sub-meters *(mikondo)* 274
8.5.4. Transcending the Shairi Matrix: The 19th-Century Origins of Mashairi ya Vidato 275
1. Shairi Meter: The Stanzaic Chaining Matrix and Its Algebraic Equation 278
2. On the Etymology and Structural Logic of the Mashairi ya Vidato .. 279
3. A Classical Blueprint: Illustrating the Vidato Mechanism through Muyaka's Uwongo 279
8.6. Confronting the Formal Models with Empirical Reality 281

CHAPTER 9. EMPIRICAL PROOFS: EXPERIMENTAL VERIFICATION OF THE FORMAL SYLLABIC AND ACCENTUAL MODELS 283

9.1. Epistemological Prelude: Falsifiability vs. Ptolemaic Epicycles 283
9.2. Protocol of Experimental Reproduction: The Three-Step Verification 284
9.2.1. The Recorded Corpus: Reading vs. Singing (Excluding Layer 3) 285
9.2.2. Speaker Selection and Acoustic Control 285
9.2.3. Syllabic Segmentation and Accentual Mapping: The Tripartite Funnel 290
1. Syllabic segmentation and the lexical accentual structure (G-lex) .. 291
2. The auditory realization (G-aud) 292
3. Instrumental verification on a hostile sample (G-lex) 294
9.2.4. Epistemological Parameters: Defining the Conditions of Falsifiability for the Competing Metrical Models 295
1. Testing the Endogenous Syllabic Principles (Vina na Mizani): Formal Matrices and Algebraic Equations 296
Condition 1: The Absolute Mathematical Verification of the Syllabic Constants ($\boldsymbol{N}$, $\boldsymbol{M}$) 296
Condition 2: The Spatial Locking of the Algebraic Rhyme Variables ($\boldsymbol{a}$, $\boldsymbol{b}$ vs. $\boldsymbol{X}$, $\boldsymbol{Y}$) 297

Condition 3: The Syntactic Impermeability of the Metrical Caesura (+) ... 297
Condition 4: The Falsification of Topological Permutations and Stanzaic Chaining ... 298
2. Testing the Exogenous Accentual Theory: Rhythmic Axioms and Cognitive Prototypes ... 298
Condition 5: Testing the Bipartite Accentual Axiom (G=2) ... 299
Condition 6: Testing the *Grenzschema* as a Cognitive Prototype ... 300
Condition 7: Testing the Periodicity of Asymmetrical Distributions (|***Ai***|) ... 300
Condition 8: Testing the Tetrasyllabic Foot against Extreme Metrical Boundaries ... 301
9.2.5. Corpus and Selection Criteria ... 302
1. List of Selected Works for the Corpus ... 303
2. Corpus Selection Criteria ... 305
Criterion 1: Diachronic Persistence and Canonical Authority ... 305
Criterion 2: The Epistemological Necessity of the "Adverse Terrain" ... 305
Criterion 3: Typological Exhaustivity ... 305
Criterion 4: Geographic Transversality ... 305
Criterion 5: Falsification by Increasing Complexity (Extreme Boundaries) ... 306
Criterion 6: Socio-Pragmatic Plurality ... 306
Criterion 7: Heterogeneity of Sources (The Anti-Editorial Bias Shield) ... 306
9.2.6. Methodological Limitations and the Horizon of Open Science ... 307
Limitation 1: Sample Size and the Big Data Imperative ... 307
Limitation 2: Editorial Dependency and the Ajami Archival Deficit ... 307
Limitation 3: Geographic Scope and Spatial Transmission ... 308
Limitation 4: The Ethnomusicological Horizon and the Exploration of Layer 3 ... 308
Limitation 5: The Single Annotator ... 309
9.2.7. Open Science Repositories and Methodological Freedom ... 309
1. The Zenodo Data Repositories (Audio Files) ... 309
Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau) (Contemporary composition) ... 309
Fatuma Shaaban (Classical corpus) ... 310
2. The GitHub Computational Repository and Reproducibility Protocol ... 310
3. The Permanent Zenodo Archive and Heavy Assets (GitHub clone and Docker image) ... 311
4. Absolute Methodological Freedom ... 311
9.3. METHODOLOGICAL SCOPE: FORMAL READING VS. SUNG PERFORMANCE . 312
9.3.1. The Acoustic Method Illustrated: Kiswahili!! ... 313
1. How the segmentation is established ... 314

2. The line examined 314
2.a. Key to the figures 314
2.b. Figure 1 — first hemistich 315
2.c. Figure 2 — second hemistich 316
2.d. What the two hemistiches show 317
3. What this line tests 318
4. Reproduction 319
9.4. UTENZI TEXTS 319
9.4.1. Case Study I: Utenzi wa Rasi 'lGhuli — Mgeni bin Faqihi (c. 1850) *319*
1. Epistemological Context and Experimental Protocol 319
2. Methodological Caveat: The Arabic Definite Article ('l) 321
3. The Competing Formal Models 323
3.a. The Endogenous Syllabic Matrix 323
3.b. The Exogenous Accentual Model / Vierke's *Grenzschema* 324
4. Utenzi wa Rasi 'lGhuli: Exhaustive Syllabic and Accentual Mapping, Stanzas 1–40 324
4.a. Reading the entries 324
4.b. Mapping 325
4.c. Excluded lines 340
4.d. Summary 341
5. Statistical Analysis and Falsification of the Accentual Model 343
5.a. Executing Conditions 1 and 2: the syllabic matrix and the rhyme 343
5.b. Executing Condition 5: the bipartite accentual axiom (G = 2) . 344
5.c. Executing Condition 6: the *Grenzschema* as a cognitive prototype 345
5.d. Executing Condition 7: the absence of periodicity 346
5.e. Executing Condition 8: the tetrasyllabic foot 347
5.f. Analytical conclusion: the terminal anchor 348
6. Instrumental Verification of the Hostile Sample 349
6.a. Composition of the sample 349
6.b. What is measured and what is reported 350
6.c. Reproduction 351
6.d. Figures 351
Stanza 31, Line 1, Nachelea ni mjinga 351
Stanza 28, Line 4, Kutunga kwa ushairi 352
Stanza 31, Line 4, Yataka mtu mahiri 353
Stanza 32, Line 1, Wala mi si mualimu 354
Stanza 12, Line 4, Hiyo siku ya nushuri 355
Stanza 29, Line 4, Wajuao tafsiri 356
Stanza 7, Line 4, Wakuu hata saghiri 357
Stanza 22, Line 3, Usafi hata ujue 358
6.e. The concordance of the instrumental reading 358
9.4.2. Case Study II: Utendi wa Haudaji — Unknown Author (19th Century) *361*

1. Philological Context and Methodological Protocol 362
2. Identification of the Formal Syllabic Matrix 364
3. The Competing Generative Model: Vierke's Grenzschema 365
4. Utendi wa Haudaji: Exhaustive Syllabic and Accentual Mapping, Stanzas 1–40 365
4.a. Reading the entries 365
4.b. Mapping 366
4.c. Excluded lines 379
4.d. Divergences between the source text and the execution 380
4.e. Summary 382
4.f. Intra-annotator control 383
5. Statistical Analysis and Falsification of the Accentual Model 383
5.a. Executing Conditions 1 and 2: the syllabic matrix and the rhyme 384
5.b. Executing Condition 5: the bipartite accentual axiom (G = 2) . 385
5.c. Executing Condition 6: the *Grenzschema* as a cognitive prototype 386
5.d. Executing Condition 7: the absence of periodicity 387
5.e. Executing Condition 8: the tetrasyllabic foot 389
5.f. Analytical conclusion: the terminal anchor 390
6. Instrumental Verification of the Hostile Sample 390
6.a. Composition of the sample 391
6.b. What is measured and what is reported 391
6.c. Reproduction 392
6.d. Figures 393
Stanza 37, line 2, Napozewa ni Azizi 393
Stanza 21, line 3, Mno nimewatamani 394
Stanza 37, line 3, Kuwa mimi muombezi 395
Stanza 32, line 4, Kula yambo moya moya 396
Stanza 9, line 4, Yasiyomtindikia 397
Stanza 22, line 1, Usiku silali tena 398
Stanza 5, line 3, Ni zake yeye Azizi 399
Stanza 20, line 4, Siche kunikarifia 400
6.f. The concordance of the instrumental reading 401
9.4.3. Conclusion on Utenzi Meter 402
1. The syllabic model 402
2. The accentual model 403
2.a. Is the bipartite axiom verified? 403
2.b. Is the *Grenzschema* a cognitive prototype? 404
2.c. Is there a rhythm? 406
9.5. *Wimbo* Texts 409
9.5.1. Case Study III: Test of a Rupture Matrix — Wimbo wa Miti — Woman Author Unknown (c. 1800–1850) 409
1. Presentation, Dating, and Philological Context 409
2. The Rupture Matrix: The Geometric Architecture of the Wimbo 410
3. Experimental Verification 411

4. Wimbo wa Miti: Syllabic and Accentual Mapping, Stanzas 1–8 412
4.a. Reading the entries 412
4.b. Mapping 413
4.c. Excluded lines 415
4.d. Summary 416
5. Statistical Analysis and Falsification 416
5.a. What is tested here, and what is not 416
5.b. Conditions 1 and 2: the syllabic matrix and the rhyme 417
5.c. Condition 3: the status of the caesura 418
5.d. Condition 5: the bipartite accentual axiom 419
5.e. Condition 6: the *Grenzschema* 419
5.f. Condition 7: periodicity 420
5.g. Condition 8: the tetrasyllabic foot 420
9.5.2. Case Study IV: Wimbo — Test of an Asymmetrical Bicolumnar Matrix (4 + 8) — The Accentual School's Home Ground 421
1. The Epistemological Value and Limits of the Fragment 421
2. The Asymmetrical Matrix (4+8) 422
3. Verification and the Availability of the Tetrasyllabic Template 422
4. Wimbo 4 + 8: Syllabic and Accentual Mapping 424
4.a. Reading the entries 424
4.b. Mapping 424
4.c. Summary 424
5. Statistical Analysis and Falsification 425
5.a. What is tested here, and what is not 425
5.b. Conditions 1 and 2 425
5.c. Condition 5: the bipartite axiom 426
5.d. Condition 6: the *Grenzschema* 426
5.e. Condition 7: periodicity 427
5.f. Condition 8: the tetrasyllabic foot 427
9.5.3. Case Study V: Wimbo — Test of a Hyper-Fractionated Matrix — Untitled, by Sheikh Ahmed Sheikh Nabhany (1927–2017) 427
1. Presentation, Dating, and Philological Context 427
2. The Hyper-Fractionated Matrix 428
3. Verification 429
4. Hyper-fractionated Wimbo: Syllabic and Accentual Mapping 430
4.a. Reading the entries 430
4.b. Mapping 430
4.c. Excluded lines 431
4.d. Summary 431
5. Statistical Analysis and Falsification 432
5.a. What is tested here, and what is not 432
5.b. Conditions 1, 2 and 3 432
5.c. The reading 433
5.d. Condition 8: the tetrasyllabic foot 434
5.e. What the syllabic model accounts for on the same corpus 434
9.5.4. Instrumental Verification of the Hostile Sample 435

1. Why the sample is reasoned and not drawn 435
2. Why the short hemistiches are excluded 435
3. Composition of the sample 436
4. What is measured and what is reported 437
5. Figures 438
Figure 1. *Wimbo wa Miti*, stanza 8, line 2, second hemistich: *nimeziye kutambuwa* 438
Figure 2. *Wimbo wa Miti*, stanza 6, line 1, second hemistich: *hatuchi twapo uwawa*. 439
Figure 3. *Wimbo wa Miti*, stanza 5, line 1, second hemistich: *wasiyesa kani zao* 440
Figure 4. *Wimbo wa Miti*, stanza 2, line 3, second hemistich: *kunyamaa ni ujura*. 441
Figure 5. *Hyper-fractionated Wimbo*, stanza 1, line 2, first hemistich: *Kusubiri tena* 442
6. The concordance of the instrumental reading 442
7. What this establishes about the two models 444
8. Reproduction 444
9.5.5. Conclusion on Wimbo Meter 444
1. The syllabic model 445
2. The accentual model 446
2.a. Where the model can be tested at all 446
2.b. Is the bipartite axiom verified? 446
2.c. Is the *Grenzschema* a cognitive prototype? 446
2.d. Is there a rhythm? 447
2.e. Condition 8, and why this meter is decisive 447
2.f. What the twenty-two segments are 448
2.g. Conclusion on this meter 449
9.6. *Shairi* Texts: The Ultimate Complexity Tests 450
9.6.1. Case Study VI: Test of a Five-Line (6 + 9) Shairi — Waji Waji (Dua ya Kuombea Mvua), by Sheikh Muhyddin bin Sheikh bin Jahathwan Al-Waily (1798–1870) 450
1. Identification of the Formal Syllabic Matrix 451
2. Waji Waji: Syllabic and Accentual Mapping, Stanzas 1–14 452
2.a. Mapping 452
2.b. Excluded lines 457
2.c. Divergences between the source text and the execution 457
2.d. Summary 458
3. Statistical Analysis and Falsification 459
4.a. What is tested here, and what is not 459
4.b. Conditions 1, 2 and 4 459
4.c. Condition 5: the bipartite axiom 460
4.d. Condition 6: the *Grenzschema* 461
4.e. Condition 7: periodicity 462
4.f. Condition 8: the tetrasyllabic foot 462
4.g. What the corpus establishes 463

9.6.2. Case Study VII: Test of Continuity of the Classical 16-Syllable "Kupindua Kina Kimoja" Matrix — Kiswahili!!, by Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau) (2003) *464*
1. Identification of the Formal Syllabic Matrix 465
2. Exhaustive Syllabic and Accentual Mapping 466
3. Kiswahili!!: Syllabic and Accentual Mapping, Stanzas 1–18 467
3.a. Divergences between the written text, the reading and the edition 472
4. Statistical Analysis and Falsification 473
5. Conclusion on Kiswahili!! 475
9.6.3. Case Study VIII: Test of a Mashairi ya Vidato Matrix — Uwongo, by Muyaka bin Haji al-Ghassaniy (1776–1840) *476*
1. Identification of the Formal Syllabic Matrix 477
2. Uwongo: Syllabic and Accentual Mapping, Stanzas 1–4 478
3. Statistical Analysis and Falsification 479
3.a. Conditions 1, 2, 3 and 4: the syllabic constant, the caesura, and the stanzaic chain 479
3.b. Condition 8: the tetrasyllabic foot, arithmetically 481
3.c. Condition 5: the bipartite accentual axiom 481
3.d. Condition 6: the *Grenzschema* 481
3.e. Condition 7: periodicity 482
3.f. Inter-pass control 482
3.g. What this corpus establishes 482
4. Conclusion on Uwongo 483
9.6.4. Instrumental Verification of the Hostile Sample: Waji Waji, Kiswahili!!, and Uwongo Combined *485*
1. Composition of the sample 485
2. Figures 487
Figure 1. *Waji Waji*, stanza 1, line 5, second hemistich: *kula siku lisilokoma*. 487
Figure 2. *Kiswahili!!*, stanza 2, line 4, first hemistich: *kosa langu kosa gani*. 488
Figure 3. *Uwongo*, stanza 1, line 3, second hemistich: *kulla jambo kwa ndiaye*. 489
Figure 4. *Waji Waji*, stanza 1, line 3, first hemistich: *warehemu wana*. 490
Figure 5. *Kiswahili!!*, stanza 1, line 4, second hemistich: *mbona mwanipija zitha*. 491
Figure 6. *Uwongo*, stanza 1, line 1, second hemistich: *leo unitumikiye*. 492
Figure 7 (random draw). *Waji Waji*, stanza 4, line 2, second hemistich: *ufishao na kufufuwa*. 493
Figure 8 (random draw). *Kiswahili!!*, stanza 3, line 3, first hemistich: *nizee wanasiasa*. 494

Figure 9 (*Grenzschema*). *Waji Waji*, stanza 1, line 4, second hemistich: *hali yetu kati chaka.* ... 495
Figure 10 (*Grenzschema*). *Kiswahili!!*, stanza 1, line 1, first hemistich: *kunyamaa nimechoka.* ... 496
Figure 11 (*Grenzschema*). *Uwongo*, stanza 1, line 1, first hemistich: *karatasi nakutuma.* ... 497
3. What the sample establishes ... 497
3.a. The triple convergence: an axiom destroyed without appeal to instrumentation ... 498
3.b. The terminal anchor exposed: *Kiswahili!!* 2.4, H1 ... 499
3.c. The internal beat manufactured: *Kiswahili!!* 3.3, H1 ... 500
3.d. The consolidated pattern ... 501
3.e. What this establishes ... 502
9.6.5. Conclusion on Shairi Meter ... 502
1. The syllabic model ... 503
2. The accentual model ... 504
2.a. Is the bipartite axiom verified? ... 504
2.b. Is the *Grenzschema* a cognitive prototype? ... 504
2.c. Is there a rhythm? ... 505
2.d. Is the tetrasyllabic template used? ... 506
3. Two independent controls ... 506
4. The instrumental sample ... 506
5. What the shairi meter establishes ... 507
9.7. The Prose Control Corpus ... 508
9.7.1. Why This Control Exists ... 508
9.7.2. How the Corpus Was Built ... 509
9.7.3. Mapping ... 510
9.7.4. Results ... 516
9.8. The Marked-Syntax Control ... 518
9.8.1. Why This Second Control Exists ... 518
9.8.2. Mapping ... 518
9.8.3. Results ... 520
9.9. Conclusion: What the Three Meters Establish ... 522
9.9.1. The Syllabic Model, Condition by Condition ... 522
9.9.2. The Accentual Model, Condition by Condition ... 523
9.9.3. Where the Confusion Is Located ... 527

CHAPTER 10. METAPOETIC FALSIFICATION: A POEM ENGINEERED TO SATISFY THE ACCENTUAL MODEL ... 529

10.1. Why a Generative Claim Must Be Constructible ... 529
10.2. *Utenzi wa Wasomi na Washairi* (The Poem of the Scholars and the Poets) — Mawazoni (2026) ... 530
10.3. Metrical Verification ... 531
10.4. Consequences for the Metrical Analysis of This Book ... 532

CHAPTER 11. GENERAL CONCLUSION: THE DISMANTLING OF THE PROCRUSTEAN BED AND THE CALL FOR OPEN SCIENCE 535

APPENDICES 548

A.1. CLASSICAL SWAHILI CORPUS: SOURCE TEXTS PROVIDED TO FATUMA SHAABAN 549

1. Utenzi wa Rasi 'Lghuli 549

2. Utendi wa Haudaji 554

3. Nyimbo 560

3.a. Wimbo wa miti 560

3.b. Sina Hali 561

3.c. Kukwepuka nana 561

4. Waji Waji. Dua ya Kuombea Mvua 561

5. Uwongo 564

A.2. CONTEMPORARY CORPUS: KISWAHILI!! (VOICED BY THE AUTHOR, USTADH MAU) 565

A.3. PROSE CONTROL READING TEXT: SOURCE MATERIAL PROVIDED TO FATUMA SHAABAN 567

A.4. MARKED-SYNTAX CONTROL CORPUS 569

REFERENCES 571

METHODOLOGICAL NOTE ON COMPUTATIONAL ASSISTANCE 579

Introduction: The Root Misconception of Exogenous Accentual Metrics

In the history of science, when a theoretical model contradicts empirical reality, the researcher is faced with a choice: alter the model or alter the data. In ancient Greek astronomy, when the observed movements of the planets contradicted the philosophical model of perfect circular orbits, astronomers invented epicycles—complex, ad hoc geometric rules—to force the data to fit the theory. This epistemological maneuver is known as σῴζειν τὰ φαινόμενα (*sôzein ta phainomena*, [saving the appearances]) (Duhem, 1908/1969).

In the history of Swahili literary studies, a strikingly similar methodological error occurred: the attempt to force an endogenous, syllabic system of metrics onto a Procrustean bed of imported accentual metrics. This epistemological blindness took two distinct paths, unfolding chronologically over more than a century, within one line of Western scholarship.

The first, and historically most pervasive, is the Eurocentric projection. An early and influential instance of this bias appears in 1870 in the work of the British missionary and linguist Bishop Edward Steere. In the preface to his *Swahili Tales* (Steere, 1870), Steere describes Swahili verse primarily through its sung performance:

> All these verses are intended to be sung, not read [...] The first impression which all negro singing makes is that it is a mere discordant jangle; but when the ear is accustomed to it, it is found to be music [...] Swahili verse is generally marked by a sort of anapæstic accent (Steere, 1870, p. xi)

Setting aside the colonial tone of the description, this passage reveals a double misunderstanding. First, Steere fails to perceive the strict syllabic architecture of the written poetic text, focusing instead on its auditory surface as performed. Second, he conflates two distinct analytical levels: the metrical structure of the verse and its musical realization in performance. His initial characterization of the singing as a "discordant jangle" underscores the perceptual gap produced by unfamiliar musical norms, even if he later concedes its complexity.

However, by ultimately interpreting Swahili verse in terms of an "anapæstic accent" he projects a classical Greek metrical framework onto a tradition governed by different principles, whose true foundations are rooted in the morphosyntactic and phonological structures of Swahili, a Bantu language. The result is a misidentification both of the internal metrical structure of the text and of the relationship between poetic form and performance.

To be precise, Steere's initial mischaracterization was not an active attempt to "save the phenomena"; rather, as an early outsider encountering an unfamiliar tradition, he simply projected a familiar rhythmic framework onto what he heard. However, this foundational error was neither abandoned nor corrected upon closer examination. Instead, it became the axiomatic baseline for successive scholarship, ultimately preserved in the 21st century through a growing apparatus of theoretical epicycles even as empirical counterevidence mounted.

An early architect of German Swahili philology formalized this initial perceptive failure into a rigid academic doctrine. As early as 1894, Carl Gotthilf Büttner, in his *Anthologie aus der Suaheli-Litteratur*, explicitly suggested that Swahili verse was governed by an underlying temporal rhythm rather than an absolute syllable count. His Eurocentric doctrine nonetheless oscillated, regarding the *utenzi*'s prosody, between "iambic" and "trochaic" feet of the classical Greek poetry (Büttner, 1894, p. xv).

Büttner compounded Steere's analytical error by conflating the internal metrical structure of the poetic text with its vocal performance. However, as Clarissa Vierke remarks, Büttner's hypotheses were immediately contested in the nineteenth century on the grounds of the rhythmic variability of performance:

> Already in the 19th century, Zache (1895, p. 374) raised an objection to any rigid prosodic pattern of the kind listed above, pointing out that the reciter had freedom to vary the accentuation, but his statement seems to have passed unnoticed by subsequent researchers. (Vierke, 2011, p. 30)

The classic Swahili meters of *utenzi*, *wimbo,* and *shairi* are traditionally written to be sung, yet academic scholarship largely bypassed this textual dimension to focus exclusively on performance. This oversight may well have suited the colonial imperative of bringing civilization to East Africa; at a deeper ideological level, acknowledging an ancient, endogenous written culture risked

unsettling the prevailing paradigm—much like the long-standing colonial refusal to attribute the stone structures of Great Zimbabwe to local African agency.

The principles of syllabic organization at work in written texts went entirely unnoticed, and the textual dimension was left unanalyzed. Meanwhile, based on oral performance, certain theorists claimed to identify accentual metrical principles (Steere, 1870; Büttner, 1894), while others—still relying on oral variability—argued that no rigid organization existed (Zache, 1895). Confusion reigned entirely.

Throughout the nineteenth century, scholarship failed to account for the key distinction between the graphic reason at work in written Swahili manuscripts and oral performances—whether chanted, cantillated, or sung—whose actual realizations superimpose the variability of different possible rhythms (*mahadhi*) upon the texts' spatial organization and graphical foundation. While Zache's observation was correct, fully understanding Swahili metrical principles still required a clear distinction between the metrical framework of the texts and the inherent indeterminacy of performance.

This confusion persisted into the early twentieth century, cementing the accentual school of classical Swahili poetry, whose primary objective remained the pursuit of rhythmic rules within performance alone. Yet, as time progressed, the distinction between the principles of syllabic organization operating within the graphical grid of written texts and oral performance became increasingly evident. Rather than adjusting to this reality, the accentual school crystallized an initial error into a dogmatic stance, deploying Ptolemaic epicycles to "save the phenomena" of perceived accentual rhythms within performance.

Alice Werner acted as the vital conduit for this bias into the English-speaking academic world. Her assertion that "[...] the verse is measured by beats rather than by syllables [...]" (Werner, 1917-20, p. 116, as cited in Vierke, 2011, p. 28) codified the error for a broader Western audience. She laid the foundations of the modern accentual paradigm in her 1917–1920 studies, arguing that Swahili verse is measured by a framework of "two beats in every line" (Werner, 1917-20, p. 116, as cited in Vierke, 2011, p. 29) rather than by syllables. While noting that one beat consistently falls on the penultimate syllable of a line due to Swahili phonology (e.g., the 7th syllable in a classical octosyllabic *utenzi* line), she observed a flexible accentuation in the earlier positions across performances for the other beat.

However, by remaining anchored in the fluidity of oral recitation, this tradition bypassed the underlying syllabic organization inscribed in the texts.

The key distinction that previous scholarship missed is that a written line carries, in the words the composer selected, a fully determined accentual structure: Swahili assigns one lexical stress to each phonological word, on its penultimate syllable, and the number and position of those stresses follow from the lexical choices alone. That structure is a property of the words the text contains: of the language, not of the meter. What a performance does with it is another matter entirely: in performance a single line can accommodate multiple different stress patterns depending on the chosen melodic realization *(mahadhi)*, including patterns that are not possible in the normal accentuation of Swahili, altering both the total number and the specific locations of the prominences.

We must therefore acknowledge that, within the domain of performance, Alice Werner's theory may well hold true—meaning that certain melodies can indeed impose strictly two stresses (considered as "beats" in Werner's view) per line of sung verse throughout an entire recitation. Conversely, other melodies may impose a greater number of stresses (three, four etc.) and/or asymmetrical stress patterns (shifting from two stresses in a line to four in the other), hence falsifying Werner's theory. All of this belongs to the realm of ethnomusicology and can be observed through acoustic analysis waveforms using audio analysis software, which also reveals numerous variations in stress patterns resulting from different melodies.

To avoid falling into the trap of performance indeterminacy, we will systematically analyze poems delivered through normal, stable readings by two L1 native Swahili speakers (a single literate reader voicing the classical corpus, and a contemporary poet reading his own work). While Edward Steere rightly stated that "[...] these verses are intended to be sung, not read" (Steere, 1870, p. xi), he omitted a key prerequisite: these poems are fully composed before they are sung. When the composer (*mtunzi*) of classical Swahili poetry composes, he builds each line on a strict syllabic grid, whether he resolves it mentally, as Shariff describes (1988, p. 70), or on the page, which fixes what he composed. This process is governed by graphic reason: the physical inscription of the poem objectifies the verse, transforming it from a fleeting temporal utterance into a stable visual and spatial architecture. Within this graphic space, the metric boundaries—

defined by exact syllable counts (*mizani*) and rhymes (*vina*)—act as rigid geometric edges that arrest the flow of speech and lock the text into an invariable arithmetical balance. As the master poets of the Swahili language themselves explicitly describe their craft, this graphic reason is exclusively syllable-based (see Chapters 4 & 5). The composers operate as textual architects manipulating statements of equality on the page, not as musicians in search of temporal "accentual beats" or imported Western and Arabic rhythmic "feet". The grid itself is verified on the text; whatever acoustic analysis the accentual claims require must therefore be conducted on read versions of the poems, rather than on sung performances.

This approach allows us to determine whether specific stress patterns operate as an intrinsic component of the texts' internal metrical structure. This is precisely where genuine scientific inquiry begins, standing in stark contrast to the Procrustean bed of the accentual school—a reference to the mythical Greek rogue who forcibly stretched or chopped off his victims' legs to fit them to his iron bed—as it selectively retains only those performances that confirm its theory while discarding all others through a process of cherry-picking.

The reach of this epistemological distortion was so pervasive that it eventually influenced even the most prominent figures of 20th-century linguistics. Joseph Greenberg (1947), despite his monumental contributions to African linguistics, inherited the stress-based assumption (Vierke, 2011, p. 31). While analyzing Swahili prosody, Greenberg conceded that the metrics should be accentual, yet admitted the system felt vague and difficult to systematically pin down—a natural consequence of searching for a rhythmic pulse that simply does not exist in the formal reality of the manuscripts.

The insistence on projecting exogenous frameworks continued unabated into the late 20th century. Even a titan of Swahili philology like J.W.T. Allen (Allen, 1971, p. 25, as cited in Vierke, 2011, p. 29) fell victim to this illusion, attempting to map classical Greek "ionic" feet onto Swahili verse.

Contemporary scholars, notably Clarissa Vierke (2011), have perpetuated and complexified Alice Werner's error. In her stylistic analysis of the *Utendi wa Haudaji*, Vierke explicitly resurrects Werner's ghost, placing Werner's quote at the very opening of her section on metrics to validate her own approach:

> "[...], and the conclusion I am led to adopt is that the verse is measured by beats rather than by syllables [...]" (Werner, 1917-20, p. 116, as cited in Vierke, 2011, p. 28)

To complexify and defend this absolute fidelity to Werner's two-beat axiom, Vierke must reconcile the theory with the erratic accentual reality of the classical manuscripts. To rescue Werner's paradigm, she introduces an elaborate epistemological epicycle: the *Grenzschema*, a concept borrowed from Roman Jakobson. Vierke posits the *Grenzschema* as an abstract cognitive prototype—specifically, a perfect, symmetrical four-syllable bisection (4+4) with accentual beats anchored on the third and seventh syllables—which the poet allegedly held in their mind as an ideal rhythm. However, Vierke herself concedes that this ideal prosodic scheme is an abstraction that is "hardly ever completely realised" (Vierke, 2011, p. 37) in the actual text. By elevating a lexical coincidence, realized in a minority of lines, into an a priori generative law, she insulates the accentual theory from empirical reality, relying on an unobservable mental intention to justify a rhythm that fails objective verification and is unknown to the poets themselves.

When the Swahili text naturally defies these regular accentual beats, Vierke invents ad hoc exceptions—claiming an accent covers two syllables here, or three syllables there, or relying on abstract rhythmic groups—simply to ensure the data eventually fits the imported model. While she establishes the philological text of the *Utendi wa Haudaji*, she completely fails to map or present these hypothesized accents systematically across the manuscript. To justify this phantom rhythm, Vierke explicitly excludes the normal reading of the poems, shifting her analytical foundation entirely to the highly variable realm of oral performance—a domain of melodic interpretation and acoustic fluctuation that this study will later formally define as "Layer 3" (see section 3.3).

She attempts to isolate a specific performative mode by drawing a distinction between *kuimba* (singing, which she admits imposes its own rhythm and overrides speech intonation) and *kukangaya* (chanting or scanning). However, this justification rests on a profound lexicographical and cultural contradiction. Vierke openly admits that she "could not find *kukangaya* [...] in any of the Swahili dictionaries" (Vierke, 2011, p. 24, footnote 1). Yet, in the very same footnote, she summarily dismisses the term *kughani*, erroneously claiming it refers exclusively to Arabic Quranic recitation.

In reality, *kughani* (to recite, chant, or sing poetry) and its agent noun *mghani* (the performer), pl. *waghani,* are the universally accepted Swahili terms utilized by the practitioners themselves, a fact codified across the entire spectrum of standard reference works. In his authoritative *Kamusi ya Ushairi* [Dictionary of Poetry], K.W. Wamitila explicitly defines *ghani* as "-imba [to sing]" or "-imba shairi [to sing poetry]" (Wamitila, 2006, p. 46). The standard *Kamusi ya Kiswahili Sanifu* (TUKI, 2nd ed., 2004) corroborates this by simply cross-referencing it to "imba [to sing]". Most devastating to Vierke's claim is the *Kamusi Kuu ya Kiswahili*, the official dictionary published by the Tanzanian National Swahili Council (Baraza la Kiswahili la Taifa [BAKITA], 3rd ed., 2022), which defines the verb *ghani* as "imba kwa sauti ya kishairi" (to sing with a poetic voice) and provides the explicit usage example: "Wanafunzi walighani mashairi kwa ustadi mkubwa" (The students recited/sang the poems with great skill). Providing the ultimate historical refutation of Vierke's claim, Charles Sacleux's foundational 1939 dictionary had already explicitly traced the term to its Arabic root (غني, *ghani*), defining it precisely as "chanter, décrire ou chanter qqn dans ses vers [to sing, to describe or sing someone in verse]" (Sacleux, 1939, p. 246).

By contesting the standard, universally documented endogenous terminology in favor of an obscure, undocumented word, Vierke artificially carves out a highly specific, yet entirely nebulous and unverifiable performative mode just to "save the phenomena" of her accentual theory. If her theory of exactly two accentual feet per line were a genuine metrical law, rigorous methodology, by her own standard, would dictate having a performer chant the epic's hundreds of lines in their entirety, followed by a systematic mapping of every accent to prove a regular statistical distribution. She does no such thing. Instead, we are left with no means of verification, as she provides accentual markings for only an isolated line or two that conveniently happen to align with her preconceived two-beat model. By cherry-picking these rare performative instances, she completely bypasses the deterministic, syllabic reality of the written grid. As will be formally established later (see section 3.3), this methodological error consists of anchoring the analysis in the non-deterministic acoustic fluctuations of oral performance (Layer 3), rather than grounding it in the invariant architecture of the text itself (Layers 1 and 2). Consequently, the structural blueprint is verified on the text, and the accentual claims are tested on normal, stable readings, which strip away the melodic distortions of performance.

More recently, a second, more subtle path of distortion has fully crystallized in contemporary comparative studies: the Arabo-centric diffusionist projection. Exemplified by recent works such as Emiliano Minerba's *Meters in Motion* (2025), this approach formally concedes the syllabic reality of classical Swahili metrical principles but attempts to rescue exogenous accentual theories through historical derivation. Relying on an argument from authority, Minerba uncritically absorbs Clarissa Vierke's unverified *Grenzschema* to serve his own confirmation bias. He observes genuine structural isomorphisms between classical Swahili stanzas and medieval Yemeni *ḥumaynī* poetry, arguing that complex rhyme patterns were borrowed across the Indian Ocean. However, he commits a profound typological error—a "homonymic illusion"—by conflating the borrowing of a spatial, architectural shell (a stanzaic layout) with the importation of a deep metrical engine. By insisting that these Swahili forms evolved through a "creolization" with Arabic tetrasyllabic feet to produce a hypothetical yet unverified "trochaic innovation", this evolutionary method ultimately strips the Swahili tradition of its endogenous formal sovereignty. Whether the distortion stems from the first path of Eurocentric projection—superimposing classical Greek metrical feet—or from this second path of Arabo-centric diffusionism, which clings to the lingering echoes of a phantom historical rhythm, the epistemological failure is identical. As their authors state them, both lineages are shielded from falsification by epicycles; once reduced to testable conditions, as this book reduces them, both are falsified. They operate as Ptolemaic epicycles, forcing an autonomous, syllabic African architecture onto a Procrustean bed of imported templates, rather than allowing the raw, observable data to naturally generate the metrical model.

To break free from this century-long epistemological impasse, a radical methodological rupture is required. The era of subjective literary hermeneutics and abstract academic speculation must finally yield to the uncompromising demands of the empirical sciences. This manifesto does not seek to counter the accentual school by merely offering another competing rhetorical interpretation, nor does it engage in endless, circular philological debates over aesthetic preferences. Rhetorical argumentation for the sake of argumentation is precisely the mechanism that has allowed exogenous paradigms to survive for so long, insulated from the friction of material reality. In this inquiry, abstract rhetoric is entirely set aside: the ultimate, absolute, and only sovereign arbiter is the raw data.

Because the Swahili meter manifests as countable syllables and rhymes on the page, and because the claims made against it concern measurable events in speech, both are open to objective scientific observation. There is no epistemological justification for hiding behind impenetrable jargon, abstract "rhythmic tendencies" or untestable cognitive prototypes when the tools of acoustic analysis, statistical mapping, and mathematical formalization are readily available. Too often, academic obfuscation has been deployed as an analytical shield: by leveraging hyper-specialized, highly intricate theories derived from foreign poetic traditions—most notably complex classical Arabic prosody—theorists erect a wall of dense terminology that discourages verification and forces exogenous models onto Swahili poetry. This reliance on theoretical obfuscation replaces scientific proof with academic posture. However, when a framework posits the existence of a metrical beat or a structural rhythm, that phenomenon must be observable in the stresses the text's own words produce and in their realization in reading, beyond what ordinary prose produces, or it must be irrevocably discarded.

Consequently, this book undertakes a massive, systematic enterprise of scientific verification grounded in a strictly Popperian methodology of falsifiability (Popper, 1959). To achieve this, we must transition from impressionistic literary descriptions to predictive mathematical models. The first fundamental step of this solution is the rigorous formalization of the endogenous metrical system.

By listening directly to the master poets of the Swahili language, who wrote about the rules of their own art, and by extracting these principles directly from the rigorous empirical analysis of the manuscript texts themselves, we construct models that are data-driven. This process initiates a continuous, iterative feedback loop between raw data and theoretical abstraction: we derive the formal models from the empirical data, translate the architectural laws of *vina* (rhyming chains) and *mizani* (syllabic count) into precise algebraic equations and spatial matrices, and subsequently subject these models to rigorous experimental testing against the texts, refining them through successive iterations. This ensures that the Swahili metrical grid is mapped exactly as it was engineered by its creators: as an uncompromising, geometric, and arithmetic reality grounded entirely in the material evidence of the corpus.

However, true scientific rigor demands that we do not merely assert the superiority of the endogenous model, but that we subject the competing theories to the exact same empirical crucible. Therefore,

the second crucial step is the strict mathematical formalization of the imported accentual positions. The elusive rhythmic beats, the "mixed-quantifying" systems of Clarissa Vierke, and the diffusionist "trochaic innovation" of Emiliano Minerba will be stripped of their interpretive ambiguity. Their paradigms will be reduced to strict numerical parameters, establishing the exact spatial and statistical conditions required for their hypothesized accentual feet and *Grenzschema* to be considered scientifically valid.

Once both the endogenous syllabic models and the exogenous accentual models are formalized, they will be confronted with the unyielding reality of the classical manuscripts. The confrontation proceeds in three steps, set out in section 9.2.3: the accentual structure the lexicon assigns to every line of the corpus, derived by rule from the printed text; the realization of that structure in a formal reading by our two L1 speakers, established by systematic listening and expressly excluding the melodic distortions of sung performance; and, offered for information and never as a source of the meter, instrumental measurement in Praat on a hostile sample: lines chosen to be the hardest for the accentual model, lines drawn at random, and lines realizing its ideal configuration. Every measure is also taken on a control corpus of ordinary Swahili prose, read under the same protocol, so that what poetry shares with ordinary speech is not mistaken for meter. By cross-referencing these datasets with large-scale statistical analysis of the texts, we will measure which theoretical model actually survives contact with the data.

This vast verification enterprise aims to close the chapter on the accentual illusion. By showing that the exogenous theories are falsified by the corpus itself, measured against ordinary prose, this work establishes that the classical Swahili metric system requires no imported metrical prosthetics. The data will ultimately reveal what the practitioners have always known: Swahili classical poetry is a hyper-sophisticated, sui generis architecture, governed by its own rhyme and its own count, *vina na mizani*.

Ultimately, an inescapable question remains: how could an entire academic lineage cite the poets' own account of the rules of their craft, and then look past it, presuming instead that these masters were unconsciously executing a phantom rhythmic design that they never named, never taught, and never knew?

Quick Reference Guide: Matrices and Algebraic Expressions of Swahili Classical Meters

Note to the reader: This guide is placed here as a tool for quick consultation during the reading of the formal and empirical chapters (Chapters 8 through 10) as well as a quick reference for readers seeking an immediate overview of the Swahili classical meters. It is not necessary to master these matrices and algebraic equations to read Chapters 1 through 7.

QRG.1. The Symbolic Alphabet

$\boldsymbol{L}$**:** Represents a discrete formal line of verse *(mshororo)*.

$\boldsymbol{L_n}$**:** Represents the sequential rank or position of the line within a stanza (*ubeti*).

$\boldsymbol{n_{max}}$**:** The parameter representing the upper bound of the possible values for L_n; its value defines the meter's formal limit.

$\boldsymbol{H}$**:** The number of hemistiches (*vipande*) per line L.

$[\boldsymbol{N}]$**:** The absolute number of syllables (*mizani*) within a specific segment or hemistich.

$\boldsymbol{S}$**:** The total syllable count (*mizani*) per line L. It is defined as the sum of all $[N]$ segments within the line ($S = [N_1] + [N_2] \ldots + [N_H]$).

> (hierarchical dominance between two forms of one variable): Placed between two indexed forms of the same metrical variable, this symbol records that the form on its left is the stronger condition and the one tested first. Written $\mathbf{F^s > F^v}$, it asserts two things at once. Logically, satisfying the left-hand form entails satisfying the right-hand one, while the converse does not hold: the two are nested and not alternative, and a corpus that meets the stronger condition automatically meets the weaker. Procedurally, the order of testing follows the symbol: the left-hand form is confronted with the corpus first, and the right-hand form is examined only where the left-hand one has been contradicted, the contradicting lines being reported. The symbol is therefore not decorative. It prevents a weaker form from

being declared as a convenience, since declaring it presupposes a published failure of the stronger.

a, b, c… (lowercase) - Local rhymes (*vina*): Whose scope is limited to the individual stanza, and which change from one stanza to the next, binding the lines of a single strophic unit. A matrix is written with the variable unresolved, as *a*: the model states the parameter, and the corpus determines its form.

Inductive modeling of the classical corpus establishes that a local rhyme is attested in two forms, which stand in the hierarchy:

$a^s > a^v$

a^s, the *syllabic local rhyme*, requires that the lines it binds close on the same full consonant-vowel sequence, onset and vowel alike.

a^v, the *vocalic local rhyme*, requires the identity of the terminal vowel alone, the syllabic onset being free to vary: closings on *-a*, *-ya* and *-wa* satisfy it, since all three end on the same vowel.

The two are not alternative conventions on the same level. a^s is the rule of the tradition. a^v is a license, taken where the lexicon offers no syllabically rhyming word of the required meaning and measure, and it is by its rarity that it is recognized as one: a form employed in a handful of stanzas out of forty is a derogation from a rule, not a second rule.

Because a local rhyme is renewed at each stanza, the license may be taken in one stanza and not in the next, and the result is therefore reported as a count of stanzas. A corpus is described as employing a^s in so many stanzas and a^v in so many, and the second figure measures the lexical pressure the composer met, stanza by stanza.

Each form is falsifiable on its own terms. A stanza declared a^s is contradicted by any bound line departing from the full syllable; one declared a^v is contradicted by any bound line departing from the vowel. Because the order of testing is fixed by the symbol, a stanza is never declared a^v without the failure of a^s having been established and the contradicting lines named.

X, Y, Z… (uppercase) - Global rhymes (*vina*): Acting as macro-structural constants, they anchor specific structural boundaries — the terminal rhyme *bahari*, for instance — across every stanza throughout the entire poem. A matrix is written with the constant unresolved, as X: the model states the parameter, and the corpus determines its form.

Inductive modeling of the classical corpus establishes that a global constant is attested in two forms, which stand in the hierarchy:

$X^s > X^v$

X^{s}, the *syllabic global rhyme*, requires the identity of the full consonant-vowel sequence at the anchored boundary: every anchored line of the poem closes on the same syllable, onset and vowel alike.

X^{v}, the *vocalic global rhyme*, requires the identity of the terminal vowel alone, the syllabic onset being free to vary.

The two are not alternative conventions on the same level. X^{s} is the rule of the tradition. X^{v} is a license, taken where the lexicon offers no syllabically rhyming word of the required meaning and measure, and it is by its rarity that it is recognized as one.

Because a global constant governs every anchored line of the poem at once, a single license anywhere resolves the constant for the whole work, and the result is therefore reported for the corpus as a whole rather than stanza by stanza.

Each form is falsifiable on its own terms. A corpus declared X^{s} is contradicted by any anchored line departing from the full syllable; one declared X^{v} is contradicted by any anchored line departing from the vowel. Because the order of testing is fixed by the symbol, a corpus is never declared X^{v} without the failure of X^{s} having been established and the contradicting lines named.

| (Vertical bar): The disjunction operator. It indicates a structural alternative or a mutually exclusive choice at a specific metric boundary.

An expression such as $([8, a] + [8, b]) \mid ([8, b] + [8, X])$ signifies that the poetic grammar allows the position to be legally instantiated by either the line $([8, a] + [8, b])$ OR the line $([8, b] + [8, X])$.

+: The abstract formal marker for the metric caesura separating hemistiches (*vipande*) within a single line.

//: The abstract formal marker for the external metric boundary, indicating the end of a full line (*mshororo*) and the structural transition to the following line within the stanza.

· (Central dot): The formal concatenation operator. It denotes the strict sequential chaining of structural units to generate a higher-level form. Its specific scope is relative to the metrical model being formalized: within the syllabic model, it concatenates independent lines to construct the complete stanza.

QRG.2. Utenzi Meter

QRG.2.1. Standard Matrix and Algebraic Equation for One Stanza (*ubeti*)

To isolate the generative engine of the *utenzi*, we first present its abstract formal model, standardly formatted here according to Western typographical conventions (vertical stacking, left-to-right).

Formal Parameters:
$n_{max} = 4$, $H = 1$, $N = S = 6$ or 8 or 11

The Matrix:
L_1: [N, a] //
L_2: [N, a] //
L_3: [N, a] //
L_4: [N, X] //

The Algebraic Equation:

$$\boldsymbol{Utenzi} = \mathbf{3}(\boldsymbol{N_a}//) \cdot \boldsymbol{N_X}//$$

(N represents the chosen syllabic constant for the poem: 6, 8, or 11 syllables)

QRG.3. Wimbo Meter

QRG.3.1. The Standard Bicolumnar Matrix and Algebraic Equation (Symmetrical and Asymmetrical)

The most common *wimbo* utilizes two hemistiches ($H = 2$), but the poet (*mtunzi*) is free to define the syllabic weight of each side of the caesura. They are numerous variations (*mikondo*), including the canonical 8+8 (*nane kwa nane*), but also strictly asymmetrical grids such as 4+8 (*nne kwa nane*), 8+4 (*nane kwa nne*), 6+4 (*sita kwa nne*), and 8+5 (*nane kwa tano*).

Formal Parameters:
$n_{max} = 3$, $H = 2$, $S = N + M$ (allowing $N = M$ or $N \neq M$).

The Matrix:
L_1: [N, a] + [M, b] //
L_2: [N, a] + [M, b] //
L_3: [N, a] + [M, b] //

The Standard Algebraic Equation for the Bicolumnar *Wimbo*:

$$\boldsymbol{Wimbo_{bicolumnar} = 3(N_a + M_b//)}$$

Note: Whether the measure is an even 6+6 or a highly uneven 8+5, the algebraic generator remains perfectly constant, locking the rhymes into an immutable spatial grid.

QRG.3.2. The Rupture Matrix: Vertical Asymmetry

The absolute sovereignty of the syllable count is demonstrated when the poet actively shifts the parameters of the grid within the same stanza. There is a specific *mkondo* where the first two lines operate on a 2+8 or 4+8 grid, but the final line dramatically expands to a symmetrical 8+8 grid.

Formal Parameters (e.g., the 2+8 or 4+8 shifting to 8+8 model):
$n_{max} = 3$, $H = 2$, $S = N + 8$ and $N = 2$ or 4

The General Matrix:
L_1: [N, a] + [8, b] //
L_2: [N, a] + [8, b] //
L_3: [8, a] + [8, b] //

The Algebraic Equation for the Rupture *Wimbo*:

$$\boldsymbol{Wimbo_{rupture} = 2(N_a + 8_b//) \cdot (8_a + 8_b//)}$$

QRG.3.3. The Monocolumnar Matrix (*Wimbo wa Zina Zitoto*)

This variation is composed of a single, undivided 8-syllable block per line ($H = 1, S = 8$). In this *mkondo*, the internal caesura vanishes entirely, and the rhyme occurs exclusively at the terminal boundary.

Formal Parameters:
$n_{max} = 3, H = 1, S = 8$

The Matrix:
L_1: [8, a] //
L_2: [8, a] //
L_3: [8, a] //

The Algebraic Equation for Zina Zitoto:

$$\boldsymbol{Wimbo_{monocolumnar}} = \mathbf{3}(\mathbf{8}_a//)$$

QRG.3.4. The Hyper-Fractionated Matrix

The written matrix grids are broken into three hemistiches (e.g., 6+6+4) and even four hemistiches per line. In this structure, the boundaries are dictated exclusively by the interlocking rhyme chains.

Note: It would be mathematically impossible to map a recurring accentual foot (like a trochee or iamb) across a single line that spasmodically shifts from a 6-syllable block to a 3-syllable block, then to a 4-syllable block, and finally to a 2-syllable block. We give an example of the most fractionated grid.

Formal Parameters for the 4-Hemistich *Wimbo*:
$n_{max} = 3, H = 4, S = 6 + 3 + 4 + 2$

The Matrix:
L_1: [6, a] + [3, b] + [4, c] + [2, d] //
L_2: [6, a] + [3, b] + [4, c] + [2, d] //
L_3: [6, a] + [3, b] + [4, c] + [2, d] //

The Algebraic Equation for the Hyper-Fractionated *Wimbo*:

$$\boldsymbol{Wimbo_{hyper-fractionated}} = \mathbf{3}(\mathbf{6}_{\boldsymbol{a}} + \mathbf{3}_{\boldsymbol{b}} + \mathbf{4}_{\boldsymbol{c}} + \mathbf{2}_{\boldsymbol{d}}//)$$

The *wimbo* grids explicitly demonstrate the division of labor inherent in classical Swahili poetical arts. The composer (*mtunzi*) resolves the formal equation on the page—whether constructing a simple $3(8_a//)$ or a highly complex hyper-fractionated matrix. Once this formal blueprint is secured, the performer (*mghani*) is granted total melodic freedom (*mahadhi*). The singer may stretch vowels, introduce syncopation, or alter the tempo, precisely because the invariant *vina na mizani* grid guarantees that the verse will not collapse formally.

QRG.4. Shairi Meter

QRG.4.1. 16-Syllable Shairi Meter: Generative Matrix for One Stanza (*ubeti*)

Formal Parameters:
$n_{max} = 4, H = 2, S = 16$ (where $S = [8] + [8]$)

The Matrix:
L_1: [8, a] + [8, b] //
L_2: [8, a] + [8, b] //
L_3: [8, a] + [8, b] //
L_4: ([8, a] + [8, b]) |
([8, b] + [8, X]) |
([8, b] + [8, a]) |
([8, X] + [8, Y]) //

The Algebraic Equation for the 16-syllable *shairi* stanza:

$$\boldsymbol{Shairi}_{\mathbf{16-syllable}} = \mathbf{3}(\mathbf{8}_{\boldsymbol{a}} + \mathbf{8}_{\boldsymbol{b}}//) \cdot \left((\mathbf{8}_{\boldsymbol{a}} + \mathbf{8}_{\boldsymbol{b}}) \mid (\mathbf{8}_{\boldsymbol{b}} + \mathbf{8}_{\boldsymbol{X}}) \mid (\mathbf{8}_{\boldsymbol{b}} + \mathbf{8}_{\boldsymbol{a}}) \mid (\mathbf{8}_{\boldsymbol{X}} + \mathbf{8}_{\boldsymbol{Y}})//\right)$$

Note: While the four-line stanza operates as a highly prominent configuration—particularly within the printed romanized corpus—it does not constitute an absolute formal limit. The classical Swahili metrical system accommodates incremental architectural expansion, as formalized in the general matrix (see QRG.4.2), which admits $n_{max} \geq 4$.

The structural alternatives in the final line (L_4) correspond to specific rhyme patterns, each explicitly designated by an endogenous Swahili term. The option $([8, a] + [8, b])$ represents *kushuka* (a descending continuation of the stanza's main rhymes); $([8, b] + [8, X])$ corresponds to *kupindua kina kimoja* (the permutation of a single rhyme); $([8, b] + [8, a])$ illustrates *kupindua* (the complete inversion of both rhymes); and finally, $([8, X] + [8, Y])$ designates *kurejelea* (the recurrence, where the final line repeats as a refrain throughout the poem).

QRG.4.2. Shairi Meter: General Generative Matrix for One Stanza (*ubeti*)

Formal Parameters:
$n_{max} \geq 4, H = 2, S = N + M$ (where $S = [N] + [M]$) (allowing either $N = M$ or $N \neq M$)

The Matrix:
L_1: [N, a] + [M, b] //
L_2: [N, a] + [M, b] //
L_3: [N, a] + [M, b] //
[...]
$L_{n_{max}}$: ([N, a] + [M, b]) |
([N, b] + [M, X]) |
([N, b] + [M, a]) |
([N, X] + [M, Y]) //

The General Algebraic Equation for the *shairi* stanza:

$$\boldsymbol{Shairi}_{\boldsymbol{general}} = (\boldsymbol{n_{max}} - \mathbf{1})(\boldsymbol{N_a} + \boldsymbol{M_b}//) \cdot \big((\boldsymbol{N_a} + \boldsymbol{M_b}) \mid (\boldsymbol{N_b} + \boldsymbol{M_X}) \mid (\boldsymbol{N_b} + \boldsymbol{M_a}) \mid (\boldsymbol{N_X} + \boldsymbol{M_Y})//\big)$$

QRG.4.3. Shairi Meter: The Stanzaic Chaining Matrix (*Mashairi ya Vidato*)

In the specific chaining structure known as *Mashairi ya Vidato*, the traditional isolation of the stanza is broken to introduce a structural interlocking. This submeter traces back to the compositions of the Swahili poet Muyaka bin Haji al-Ghassaniy (1776–1840) (Hichens, 1940, pp. 92-93). The rhyme of the second hemistich of lines 1, 2, and 3 in stanza U_i permutes to the first hemistich position in the fourth line of the same stanza. This specific rhyme then functions as the first hemistich rhyme for the lines 1, 2, and 3 of the subsequent stanza U_{i+1}, sequentially chaining the stanzas across the entire poem.

To accurately model this stanzaic chaining, we abandon the static variables (a, b) used for independent stanzas and introduce an indexed mathematical sequence of variables $R = (r_0, r_1, r_2, \dots, r_k)$. For any stanza of index $i \geq 1$, the stanza inherits the first-hemistich rhyme r_{i-1} from the preceding stanza (or from an initialization variable r_0) and introduces a new second-hemistich rhyme r_i.

Formal Parameters:
$n_{max} = 4$, $H = 2$, $S = 16$ (where $S = [\ 8\ ] + [\ 8\ ]$)

The Matrix for Stanza U_i:
L_1: [8, r_{i-1}] + [8, r_i] //
L_2: [8, r_{i-1}] + [8, r_i] //
L_3: [8, r_{i-1}] + [8, r_i] //
L_4: [8, r_i] + [8, X] //

The Algebraic Equation for Stanza U_i:

$$\boldsymbol{U_i = 3\left(8_{r_{i-1}} + 8_{r_i}//\right) \cdot \left(8_{r_i} + 8_{X}//\right)}$$

The Algebraic Equation for the Complete Poem P:

$$\boldsymbol{P = \sum_{i=1}^{k} \left[3\left(8_{r_{i-1}} + 8_{r_i}//\right) \cdot \left(8_{r_i} + 8_{X}//\right)\right]}$$

To model the complete poem P of k stanzas, the equation defines a sequential concatenation over the index i. This formalization shows that the classical Swahili grid can function as a highly sophisticated generative algorithm capable of high structural complexity.

Chapter 1. Epistemological Preliminaries: Datasets and the Exo-African Postulate

1.1. The Chronological Division and the Scope of the Critique

Before proceeding, it is necessary to establish the precise scope and typological boundaries of this formal analysis. While the classical Swahili prosodic system comprises up to thirteen recognized meters (Nabhany, 2006, *Dibaji*), the primary focus of this work rests specifically on the *utenzi*, *wimbo*, and *shairi* meters. The division of the Swahili literary continuum into "archaic" and "classical" periods is a historiographical projection—a conceptual framework inherited from ancient Greek scholarly traditions and subsequently superimposed onto an African reality. Just as Hellenic literary history distinguishes the fluid, unwritten oral performances of the ἀοιδοί (*aoidoi*), epic poet-singers, in the Archaic period from the stable, written texts that crystallized following the adoption of the Greek alphabet (8th–6th centuries BC), Swahili scholarship relies on a similar, albeit *projected* and far less documented, dichotomy between archaic oral songs (like those of Fumo Liyongo) and their written codification during the classical age.

The advent of this classical period is marked by the monumental transition to a written medium utilizing the Arabic script (*ajami*). Although historians debate the exact chronological arc, this progressive textualization firmly crystallized around the mid-17th century in the Lamu archipelago. As synthesized by Vernet (2005, pp. 162-164) building upon the extensive historiographical work of Pouwels (1987, 1992, 2002), Knappert (1979), Harries (1962), Brown (1985, 1988), and J. de V. Allen (1981), this era coincided with a significant influx of sharifian and scholarly lineages from Yemen and the Hadhramaut who settled primarily in the Lamu archipelago. These migrants brought about a profound "cultural renewal" introducing a highly prestigious, orthodox Islamic literate culture. Fusing with the endogenous Bantu matrix, this intellectual catalyst

propelled the key transition of Swahili poetics from an oral tradition into a sophisticated written architecture. Specifically, as noted in the research of Knappert (1979) and Pouwels (1992), this classical age represents a conscious textualization where ancient oral literary forms were formalized and blended with Arab traditions. Ultimately, this inaugurated a golden age of manuscript production and introduced a strictly graphic rationality that permanently altered the poetic architecture.

To fully grasp the magnitude of this structural shift, it is essential to examine the historical consensus synthesized by Thomas Vernet. His doctoral thesis is a work of immense erudition, offering a massive and masterful historical synthesis of the region. Because this monumental text remains exclusively in French, an extended translated excerpt is provided here. This serves as a vital first entry point for the English-speaking readership—though engaging with the exhaustive original text directly remains highly recommended to fully appreciate its scope. In the following passage, Vernet details how the arrival of Hadhrami sharifs catalyzed this textualization, inaugurating a golden age of Swahili poetics while anchoring it in the endogenous language:

> According to Pouwels, the Hadhrami sharifs brought about above all what he terms a 'cultural renewal,' which originated in the city of Pate. It was first a religious renewal: they propagated views based more on a literal reading of sacred law and an increased dissemination of the Shafi'i legal tradition. Consequently, they contributed to remodeling the local interpretation of Islamic law and injecting it with more orthodox principles. It seems, for example, in legal matters, that they emphasized patrilineal inheritance and the alienation of land.
>
> Likewise, they contributed from Pate onwards to the expansion of a literate culture that seems to have scarcely existed before: written versification in Swahili likely developed there in the mid-17th century, based in part on structures common in the Hadhramaut. Subsequently, the 18th century became a golden age of written poetry on the northern coast—in Pate, Siyu, or Lamu—blending Arab traditions with the forms of ancient oral literature. The descendants of Abu Bakr bin Salim, in particular, formed a lineage of literate poets in Pate and later in Lamu between the mid-18th century and the end of the following century, including Sayyid Abdallah bin Ali bin Nasir (c. 1720–1820), who composed the poem *al-Inkishafi* in Pate, a

> masterpiece of Swahili poetry. Furthermore, at that time, Siyu was a major center of Islamic scholarship, and a whole industry of manuscript copying and binding flourished there. It is noted, however, that poetry or copying were not exclusively practiced by sharifs, and above all, that the unpublished works that have reached us are written in Swahili and not in Arabic. For within the span of one or two generations, the descendants of the Hadhrami migrants became Swahilized, just as the al-Mahdali had been before them. (Vernet, 2005, pp. 162–164)[1]

This historical synthesis requires two critical epistemological observations regarding the formalization of Swahili verse. First, Vernet's careful phrasing—noting that this literate culture "seems to have scarcely existed before" ("parait n'avoir guère existé auparavant")—highlights the inherent fragility of the historiographical record. As the empirical data for earlier periods is either lost or obscured, historians are forced to rely on a projection of the available data. The 17th-century classical age is not necessarily an absolute beginning *ex nihilo*, but rather the documented juncture when the ancient oral tradition crystallized onto the physical page, shifting from performance to graphic reason.

---

[1] French Original: "Pour Pouwels, les sharifs hadrami entraînent avant tout ce qu'il nomme un « renouveau culturel », qui prend naissance dans la cité de Pate. C'est d'abord un renouveau religieux : ils propagent des vues reposant davantage sur une lecture littérale de la loi sacrée et une diffusion accrue de la tradition juridique shafite. Ils contribuent par conséquent à remodeler l'interprétation locale de la loi islamique et à lui injecter des principes plus orthodoxes. Il semble par exemple, en matière juridique, qu'ils mettent l'accent sur l'héritage patrilinéaire et l'aliénation des terres. De même, ils concourent à partir de Pate à l'expansion d'une culture lettrée qui parait n'avoir guère existé auparavant : la versification écrite en swahili s'y développe vraisemblablement au milieu du XVIIe siècle, en partie fondée sur des structures courantes en Hadramaout. Par la suite le XVIIIe siècle est un âge d'or de la poésie écrite sur le rivage septentrional, à Pate, Siyu ou Lamu, mêlant des traditions arabes et les formes de la littérature orale ancienne. Les descendants d'Abu Bakr bin Salim en particulier ont constitué à Pate, puis à Lamu, une lignée de poètes lettrés entre le milieu du XVIIIe siècle et la fin du siècle suivant, dont Sayyid Abdallah bin Ali bin Nasir (vers 1720-1820) qui compose à Pate le poème al-Inkishafi, pièce maîtresse de la poésie swahili. En outre, à cette époque, Siyu est un centre majeur d'érudition islamique et toute une industrie de la copie et de la reliure s'y épanouit. On remarque cependant que la poésie ou la copie ne sont pas exclusivement pratiquées par des sharifs, et surtout que les œuvres inédites qui nous sont parvenues sont rédigées en swahili et non en arabe. Car en l'espace d'une ou de deux générations, les descendants des migrants hadrami sont swahilisés, ainsi que l'avaient été avant eux les al-Mahdali."

Second, and most for our formal approach of the classical Swahili poetry, this literary explosion was emphatically endogenous in its linguistic medium. As Vernet points out, poetry and manuscript copying were not exclusively the domain of the sharifs from Yemen—that is to say, endogenous Swahili scholars and poets also drove this literary production—and he explicitly emphasizes that "the unpublished works that have reached us are written in Swahili and not in Arabic" because the "Hadhrami migrants became Swahilized" (Vernet, 2005, p. 164). He further substantiates this in his annotations (p. 164, n. 23), observing that the oldest surviving original documents penned by coastal scholars—including early 18th-century manuscripts from Pate preserved in the Goa archives—are written strictly in Swahili using Arabic characters (*ajami*). This linguistic sovereignty proves a fundamental point: the Hadhrami influence did not replace the Swahili language with Arabic. Because the Bantu language was retained, it logically follows that the phonological rules of Swahili continued to govern the poetry. While the migrants provided the graphic technology (the script) and perhaps introduced certain structural forms—as modern comparative scholars like Minerba (2025) have noted—the poetic engine remains strictly autonomous, generated by the phonological reality of a Bantu language rather than exogenous Arabic prosody.

Throughout the 19th and 20th centuries, the tradition underwent a second major medial shift with the transcription and composition of poems in the roman alphabet. The texts mediality has naturally continued to evolve in modern times, with classical manuscript scribing largely giving way to print and broadcasting, including the internet since the end of the 20th century. To ensure maximum analytical precision, our empirical verification is primarily anchored on the mature romanized forms of the 19th century restricted to the *utenzi*, *wimbo*, and *shairi* classical meters. However, to objectively test the structural continuity of the metrical system across time, we will supplement our corpus with a specific 20th-century corpus, treating it as an additional empirical data point to observe the stability of the classical forms in contemporary practice.

1.2. The Methodological Imperative of Formal Modeling and Experimental Test

To counter a century of epistemological wandering, this work places data-driven formal metrics and experimental verification at the absolute center of its analytical framework. The method combines two movements, and it is worth distinguishing them at the outset. The formal models are built inductively: the syllabic matrices and their algebraic equations, set out in Chapter 8, were formalized from the descriptions of the Swahili-language poets and scholars and from the manuscripts themselves, and refined by successive contact with the corpus, so that a text which resisted the formalization obliged the formalization to change. The testing is deductive: once each model is stated in a form precise enough to fail, the conditions of its failure are fixed in advance and the corpus is confronted with them.

What being data-driven does not mean here is that a metrical theory can be read off an acoustic signal. No theory ever has been, and none could be, since the signal does not distinguish what a language imposes from what a meter requires. The instrument answers narrower questions, and it answers them on a sample.

The analysis accordingly begins with the printed lines. To each line of the corpus we apply the fixed penultimate stress of the Swahili phonological word — a rule of the language itself, established by a century of description and not by this study — which yields, deterministically and without any recording, the accentual structure that the composer's lexical choices produce. This dataset is the foundation of the argument, and it is the only one in this study that depends on no listener, no speaker, and no software: any reader who knows the penultimate rule recomputes it from the text as printed.

That structural dataset is then confronted with two further orders of evidence. The first is a formal reading aloud of the corpus by L1 Swahili speakers, screened line by line by ear, which records how much of the lexical accentual structure a reader actually realizes. When analyzing dialect-specific historical poetry such as classical Kiamu verse, our protocol does not require absolute dialectal matching: it deliberately leverages the macro-Swahili phonological competence of modern readers, since the penultimate stress rule shows profound synchronic and diachronic stability within the Swahili linguistic core. The reading is expressly a reading and not a performance, which isolates the analysis from the non-deterministic

musical distortions of chanting, cantillation and song, with their artificial isochrony, melismas and performative rests according to the chosen melodies (*mahadhi*).

The second is instrumental. A sample of lines, whose composition is fixed in advance and which includes lines drawn at random, is processed in Praat through the Parselmouth library to measure, syllable by syllable, fundamental frequency, duration and intensity. The sample establishes that the categories being counted have measurable physical correlates. It is a sample and not a census, and no statistic in this book rests on it.

These three datasets, kept strictly separate and published separately, allow us to strip away the subjective rhythmic projections of some philologists and to isolate the true generative engine of the Swahili poem: the invariant equation of *vina* (the syllabic rhymes) and *mizani* (the syllabic count). Consequently, formal metrics is not a sub-discipline of literary hermeneutics, nor does it concern itself with intricate layers of theoretical scaffolding that share one fundamental flaw, that of being entirely unfalsifiable. It is the exact, empirical study of the poetic architecture as it is codified on the page and as it constrains what the composer could write. This approach treats the classical written stanza not as a musical score awaiting a temporal beat, but as a sovereign graphic and formal contract.

Our approach is scientific in its philosophy: data are the core of the domain, and the primary judge of truth or falsity. Our formal matrices and algebraic equations of classical Swahili meters (*utenzi*, *wimbo*, and *shairi*) must match the data, not the reverse. Furthermore, experimental verification is mandatory within this epistemology: it requires that our formalizations be falsifiable—that is, that we be able to determine, through a clearly described experimental protocol, whether they hold true or are refuted, and to describe that protocol with enough precision for other researchers to reproduce it independently. Transparency is not optional. We accordingly avoid jargon-laden postures, and any conclusion resting on elaborate, self-immunizing argumentation while remaining only lightly anchored in data.

Grounded in both the written lines of the page and recorded readings of them, formal metrics provides the only rigorously scientific tool capable of dismantling the accentual fallacy and establishing the syllabic truth at the core of Swahili meters. This work does not claim to discover the architecture it describes: it investigates the endogenous system long articulated by the Swahili-language poets

themselves, and confirmed, with equal rigor, by the philologists who took that testimony seriously. Its primary contribution will lie in verification, rather than discovery. By subjecting this architecture to formal modeling and to experimental tests on the corpus, this study will show that the accentual hypothesis functions as a century of exogenous noise superimposed on a signal that was never in doubt: the syllabic consensus of *vina na mizani*, sounded by the poets and heard, correctly, by those who troubled to listen.

1.3. The Dammann Corpus: Historical Context, Ideological Bias, and the Neutrality of Raw Data

Clarissa Vierke's doctoral thesis (2011) relies heavily on the poem entitled *Utendi wa Haudaji* to propose her accentual model. She explicitly states that her edition is:

> [...] based on a transliteration done by Ernst Dammann and Muhamadi Kijuma during Dammann's stay on Lamu in 1936 and 1937, and two manuscripts in Arabic script from the Allen collection of the archive of the School of Oriental and African Studies in London. (Vierke, 2011, p. xiv)

The author of this *utenzi—utendi* being the Kiamu dialect form for the standard *utenzi*—is unknown, and neither the poem nor its historical context is remembered, as Vierke notes:

> [...] in Eastern Africa, not only the *"Utendi wa Haudaji"* was unknown to all the people I talked to, but also the historical incident did not ring a bell with any of the *wazee* that I talked to. (Vierke, 2011, p. 420, n. 604)

Faced with such a twilight text—whose historical memory has entirely faded from the living tradition—Dammann's role is as central to its preservation as it is deeply paradoxical. From a strictly data-driven perspective, it must be acknowledged that without Dammann's archival efforts, a vital piece of Swahili poetic heritage might have been permanently lost. Yet, the formalization of the accentual

matrices applied to it in recent literature relies on the transcription work of a "politically compromised scholar" (Poewe, 2001).

Ernst Dammann (1904–2003) was a prominent German theologian and Africanist whose linguistic fieldwork in East Africa during the 1930s yielded invaluable transliterations of classical Swahili manuscripts (Jungraithmayr, 2007). However, historical records confirm that Dammann was an early member of the NSDAP (National Socialist German Workers' Party) and served as the *Landesgruppenleiter* in East Africa during the exact period this data was collected (Poewe, 2001; Balbiani, 2023). This position was not that of a peripheral sympathizer; a *Landesgruppenleiter* was the supreme political and organizational head of the Nazi Party's foreign apparatus (the NSDAP/AO) within the territory, enforcing a legally codified racial hierarchy (Pugach, 2012).

In Sonderegger's estimation, Dammann stands as a striking embodiment of colonial-racist ideology:

> Ernst Dammann can serve as a striking example. [...] He was not only an opportunistic member of the Nazi party since 1931, but deeply entrenched in racist thought. The colonial-racist state of mind is perfectly personified by this man who, as is evidenced by his memoirs published in 1999, upheld his racist and paternalistic views until late in life. It is easy to conclude that he had a certain influence on some of his students. (Sonderegger, 2011, p. 138)

It is a fundamental principle of our methodology to strictly avoid *ad hominem* attacks and genetic fallacies. In fact, we explicitly distance ourselves from Sonderegger's concluding assumption that Dammann's ideological worldview inevitably infected his students or the modern scholars who utilize his archives. The personal moral failings or extreme political affiliations of a researcher do not automatically invalidate the empirical linguistic data they collect. Because Dammann's transcriptions can be cross-referenced with the Allen archives and show no evidence of deliberate textual corruption, the raw data—the syllables, accents and word transcribed—remains a verifiable empirical artifact. We firmly separate the validity of the data from the ideology of the compiler.

However, the epistemological issue lies in the absence of sustained critical examination of the chain of transmission and the conditions of production of this archive. Recognizing the neutrality of the raw data

does not permit treating the archive as an ahistorical entity. Vierke, working directly within this chain of transmission, accessed a sensitive archive compiled by a high-ranking Nazi official but treated the dataset purely as a decontextualized artifact, entirely omitting any critical reflection on the collector's well-documented racial ideology.

The presence of Dammann in this discussion is strictly empirical: his transliterations serve as the foundational dataset for the modern continuation and reformulation of accentual approaches by scholars such as Vierke. In this regard, modern accentual research acts as a methodological nexus, relying on Dammann's raw data while simultaneously adopting the exogenous accentual theory formulated decades earlier by scholars like Alice Werner.

It is trivially true, yet epistemologically vital to state, that Werner's theoretical framework (established in the 1910s and 1920s) predates Dammann's fieldwork (1930s). The accentual matrix is therefore historically and structurally independent of Dammann's ideology. Highlighting Dammann's political background is not an attempt to poison the well of accentual theory by association. Rather, it demonstrates a profound methodological flaw in modern studies: the uncritical treatment of colonial-era scholarly outputs as epistemically transparent foundations.

The silence here belongs entirely to modern research, which treats these archives as neutral foundations and fails to contextualize the deeply problematic colonial paradigms of their producers.

What unites these two distinct chains—the documentary and the theoretical—in modern scholarship is precisely this shared lack of contextualization. However, it is specifically along the theoretical chain that this epistemological blind spot skews the metrical analysis. By adopting Werner's exogenous theory without historical distance, modern research fails to question *why* early academia systematically struggled to recognize a sovereign, endogenous East African metrical genius in classical Swahili poetry.

The historical reasons for this early theoretical blindness are diverse and must be weighted carefully. On one level, this failure often stemmed from simple, good-faith heuristic errors: early observers, faced with an unfamiliar poetic system, naturally projected the classical Greek or Arabic metrical matrices they already mastered. At an equally important structural level, the broader colonial paradigm could provide strong incentives to downplay or externalize evidence of autonomous African cultural achievement, particularly when such

evidence challenged prevailing assumptions about European civilizational authority.

While overt scientific racism was also a prevalent force during this era, utilizing it as an analytical explanation requires explicit textual evidence for individual authors. However, the broader colonial compulsion to erase or externalize African ingenuity—the need to seek outside the Swahili sphere for the true generative mechanisms of its cultural achievements—operates as a structural epistemological barrier. This systemic reflex is an epistemic pattern we must now formally examine as an avatar of the Great Zimbabwe Syndrome.

1.4. The Great Zimbabwe Syndrome and the Exo-African Postulate

This lack of critical framing creates fertile ground for what we must term the Exo-African Postulate, a bias best understood through the historical analogy of the Great Zimbabwe Syndrome.

The true historiography of Great Zimbabwe reveals a deliberate epistemological erasure. The monumental stone ruins were credited to exogenous origins, such as King Solomon's mines, Phoenician traders, or a mysterious ancient "white race" — to anyone at all except the Shona people whose hands had laid the stone. Cecil Rhodes, whose British South Africa Company controlled the territory, actively funded compliant antiquarians, such as J. Theodore Bent in 1891, whose work kept the African builders out of the narrative (Garlake, 1973; Fontein, 2006). Professional archaeologists then showed, David Randall-MacIver in 1905 and Gertrude Caton-Thompson in 1929, that the empirical evidence clearly pointed to the endogenous Shona populations as the builders. This scientific reality collided violently with the dominant Eurocentrism of the era and threatened the ideological justification for colonial rule. The settler establishment refused to believe it for decades more.

In Swahili metrical studies, this syndrome operates not on physical stone, but on poetic form. The Exo-African Postulate is the implicit academic assumption that if an African poetic matrix is architecturally sophisticated and arithmetically rigorous, its generative engine must have been imported from outside.

When modern research—such as Vierke's 2011 analysis—utilizes the Dammann dataset, it accesses highly accurate raw data. However,

the trap of the Exo-African Postulate is sprung not by the data itself, but by the exogenous theoretical framework applied to it. By adopting the accentual theories of Alice Werner (1917-20) and applying them to the perfectly aligned "stones" of the *utenzi* while sidestepping the endogenous Swahili blueprint (the syllabic principles of *vina na mizani*), the contemporary accentual school assumes the architecture requires an exogenous explanation. The material reality of Dammann's raw text merely serves as a blank canvas upon which this paradigm legitimizes its own imported projections.

This critical misalignment is not a matter of personal intent on the part of contemporary scholars, but rather the symptom of a deep-seated disciplinary blind spot. The theoretical disqualification of the endogenous metrical rules was not initiated by Dammann, but was already deeply entrenched in the colonial paradigms of his predecessors. Viewed across a deeper historical continuum, the methodological heritage of the Exo-African Postulate reveals an enduring lineage: a foundational colonial Eurocentrism—manifested in the theoretical works of Steere (1870), Büttner (1894), and most decisively for the accentual school, Werner (1917-20)—had laid the groundwork long before Dammann compiled his corpus.

Yet, its remarkable persistence well into the first quarter of the twenty-first century rests not on scientific validity, but on an immunizing theoretical architecture. As an a priori assumption rather than a data-driven conclusion, the Exo-African Postulate will inevitably persist in Swahili poetic studies—not because classical texts empirically exhibit an accentual meter, but because the theoretical framework itself is ideologically compelled to yield an exogenous origin.

1.5. From Early Heuristics to the Analytic Turn in Swahili Metrics

Having outlined the historical context that often accompanies the projection of exogenous frameworks onto African texts, it is intellectually imperative to adopt a posture of humility. When early scholars such as Edward Steere (1870), Carl Gotthilf Büttner (1894), and later Alice Werner (1917-20) first confronted the immense, highly sophisticated corpus of classical Swahili poetry, they faced an alien acoustic and formal reality. Naturally and heuristically, they reached

for the only conceptual calipers they possessed: the classical metrics of Homer or the stress-timed rhythms of Germanic poetics. Because Arabic poetry was historically better known and more widely studied, one might be tempted to use it as a comparative framework to interpret a seemingly related poetic tradition bearing a surface resemblance, such as Swahili poetry. This is a normal cognitive reflex. Attempting to decode an unknown system by applying a familiar Arabic, Greek, or Germanic metrical grid is an honest intellectual mistake—an indispensable process of trial and error in the early stages of any discipline.

The problem arises only when these early heuristics harden into institutional dogma. For too long, the defense of the accentual model in Swahili poetry has relied on self-referential terminology, abstract "rhythmic tendencies", ad hoc hypotheses, and unverifiable concepts used to artificially sustain an aging paradigm. It has evolved into a speculative endeavor that systematically resists empirical testing.

To break this deadlock, Swahili metrical studies must undergo a rigorous analytic turn. Drawing on the core imperative of analytic philosophy—which demands strict linguistic precision, conceptual clarity, and logical rigor over opaque speculation—we must adopt an uncompromising scientific method. As Ludwig Wittgenstein stated: "What can be said at all can be said clearly; and whereof one cannot speak thereof one must be silent." (Wittgenstein, 1922, p. 27).

In our framework, this means replacing theoretical ambiguity with formal models tested against the data. Our scientific method of studying Swahili classical metrics demands that theoretical propositions be treated as collaborative research endeavors requiring clear formalization, transparent data, rigorous experimental verification and reproducibility of the experiments.

As Karl Popper insists, the strength of a scientific paradigm lies not in its initial infallibility, but in its capacity to be falsified and discarded when contradicted by experimental data (Popper, 1959). When a paradigm is falsified, the academic community must accept the rupture and evolve. Approaching this paradigm shift with proper scholarly humility, we hope that our models prove to be robustly data-driven, yet remain fully transparent and falsifiable above all else. When unverifiable theoretical constructs and self-immunizing postulates are finally abandoned in favor of strict, Popperian experimental tests and formal data processing, textual obfuscation collapses. Corpus linguistics, software-assisted audio analysis, and formal modeling applied to classical Swahili poetry mark the

definitive end of speculative dogmatism, inaugurating a truly scientific and reproducible era in Swahili metrical analysis.

In full accordance with the principles of Open Science—championing data transparency, methodological reproducibility, and open access—we make our datasets, audio files, and verification tools fully accessible. This commitment not only invites replication and critique from fellow researchers, but actively welcomes invaluable insights and suggestions, particularly from East Africa. By opening our methods to collective scrutiny, we harness the true self-correcting power of open research: errors are rapidly identified, hypotheses are rigorously tested, and our collective understanding of classical Swahili poetry progresses through collaborative, transparent scholarship.

Chapter 2. The Accentual School of Classical Swahili Poetry

2.1. Phonological Grounds for Accentual Metrics in Swahili

Before one can adequately formalize and test the theories proposed by the accentual school, it is intellectually imperative to present a transparent and objective assessment of the phonetic tools actually available to the Swahili poet (*mtunzi*). To understand how Western scholars could plausibly hypothesize the existence of an accentual rhythm within a strictly syllabic grid, we must examine the specific phonological properties of the Swahili language. By honestly outlining the phonetic reality of the language, we can determine the exact mechanical boundaries within which any rhythmic theory must operate.

The most defining prosodic feature of Swahili is a negative one: the language lacks any phonological opposition between long and short vowels, and by extension, between light and heavy syllables. Unlike Classical Arabic or Ancient Greek, where vocalic duration is phonemic and directly determines syllable weight and metric quantity, Swahili syllables are inherently isochronous in their underlying phonological structure. As linguists have conclusively established, Swahili phonology does not possess any phonological opposition between long and short syllables (Sacleux, 1909, pp. 6-10; Ashton, 1962, pp. 3-5; Polomé, 1967, pp. 36-59). A sequence of two vowels is always scanned as two distinct syllables, and there are no inherent "heavy" or "light" syllables that a poet can deliberately manipulate as modular rhythmic building blocks. This absolute lack of vocalic quantity formally precludes the massive, direct importation of quantitative metrics, rendering the exact transposition of the Arabic *ʿArūḍ* or classical Hellenic scansion impossible on Swahili soil.

However, while Swahili lacks phonemic quantity, it possesses a highly predictable and robust dynamic stress system. In standard Swahili phonology, the lexical accent is rigidly fixed: it falls overwhelmingly on the penultimate syllable of the phonological word. Exceptions to this fundamental rule are exceedingly rare, generally

confined to a limited subset of recent, unassimilated Arabic loanwords, and even these exceptions tend to rapidly conform to the Swahili penultimate pattern.

This fixed penultimate stress serves a vital cognitive and syntactic function: it acts as a definitive structural demarcator, explicitly defining the boundaries of the phonological word. Its mechanical consistency is perfectly illustrated by the morphophonological behavior of enclitics. When a stressless enclitic, such as the locative suffix *-ni*, is attached to a word (for example, the word *nyú-mba* "house", naturally stressed on the syllable *nyú*), the primary accent does not remain static. It automatically and predictably shifts one position to maintain its penultimate status in the newly formed word (*nyu-mbá-ni*, now stressed on the syllable *mbá*). It is the acoustic manifestation of this penultimate stress that provides the theoretical springboard for the accentual school. In phonetic terms, mainstream Swahili linguistics defines lexical stress primarily through an elevation in pitch (fundamental frequency, or F_0). However, this stress is also accompanied by a localized lengthening of the syllable. It is this specific property—a measurable increase in duration—that can be utilized to adapt quantitative metrical systems, such as those of classical Arabic, to Swahili. Yet, it is also this very same surface effect that the accentual school isolates to justify its own metrical frameworks. By doing so, they mistake a phonetic consequence for a generative metrical rule, overlooking the fact that Swahili establishes no phonemic opposition of syllable weight.

This stress-induced lengthening is precisely the prosodic phenomenon that permits the conceptualization of an accentual metrical rhythm. Presenting the argument of the accentual school objectively requires acknowledging this specific linguistic loophole: while one cannot build a quantitative meter out of phonemically long and short vowels, a metrician can theoretically utilize this contextual lengthening to adapt a metrical framework to Swahili phonology. By tracking these organically elongated penultimate syllables, scholars can map a predictable pattern of prominent beats, stressed syllables alternating with unstressed syllables.

This is precisely the theoretical apparatus deployed by modern proponents of the accentual school, such as Clarissa Vierke (2011) and Emiliano Minerba (2025). Deprived of true vocalic quantity, they leverage the natural acoustic prominence and the physiological lengthening of the Swahili penultimate stress to hypothesize an underlying rhythmic schema. They propose that the verse is truly

driven by these recurrent accentual peaks—where the contextual elongation of the vowel simulates a metrical foot—rather than by the sheer, flat arithmetic of the syllable count (*mizani*) and syllabic rhymes *(vina)*.

The theoretical viability of this adaptation, however, immediately collides with the necessity of empirical verification. The central epistemological question remains: are these hypothesized accentual schemas consistently realized in the oral execution of the poems? If the accentual school is valid, these stress patterns must manifest as a predictable, repeating rhythm, rather than as chaotic phonetic accidents dictated merely by word choice.

To test this reliably, one must rigorously isolate the mode of vocalization. As established previously, the classical Swahili poem is frequently performed as a song, accompanied by a superimposed traditional melody (*mahadhi*). The performative act of singing inherently subjects the text to external musical imperatives: the singer may stretch unaccented vowels to fit a prolonged musical note, insert pauses that have no textual equivalent, or syncopate the delivery for aesthetic effect. This musical freedom actively overrides, stretches, and displaces the natural lexical stress of the spoken language. By its very definition, the imposition of a free musical melody shatters the predictable phonological baseline, making the rigorous study of a strictly textual accentual metric impossible, as the necessary regularity of the accentual fields is destroyed by the performer's interpretation.

Unlike the vast multitude of interpretations permitted by traditional performance—which easily lends itself to cherry picking, retaining only the instances that confirm a preconceived theory while discarding all others—the words of the line themselves yield a single, stable accentual structure, since Swahili assigns one stress per phonological word and assigns it to a fixed position. That structure is what the composer selected, and it is recoverable from the printed text alone. A formal reading aloud does not reproduce it exactly: any reading superimposes a phrase intonation which realizes some of these nuclei and leaves others unrealized. The two must therefore be recorded separately, and they are. If statistical regularities of the kind the accentual school predicts emerge at either level, its predictive power will be empirically confirmed.

Therefore, to determine whether the accentual matrices proposed by the accentual school constitute a genuine generative law of the text, and not merely an artifact of performative singing, we must test them

first against the accentual structure the lexicon of each line contains, and then against what a formal reading of that line actually realizes. However, before proceeding to this experimental verification, we must first formalize these various accentual theories into clear, formal models.

2.2. The Accentual School: A Century of Exogenous Projection

Before delving into the contemporary and highly sophisticated models of the accentual school, it is intellectually imperative to trace the historical genesis of this epistemological divergence. In reconstructing this specific historiography, we are indebted to the erudition of Clarissa Vierke.

In her 2011 work, Vierke provides an exhaustive and masterfully documented literature review of early Western scholarship on Swahili prosody. She acknowledges that within the endogenous Swahili tradition, as explicitly confirmed by master poets like Ahmed Sheikh Nabhany and Ahmad Nassir, the classical poetic meters are primarily based on an absolute syllable count and syllabic rhyme (Vierke, 2011, p. 28). However, she catalogs immediately after this acknowledgment how generations of Western scholars—beginning in the late nineteenth century—ignored this explicit syllabic consensus to search for a hidden "rhythmicisation through accent or beat" (Vierke, 2011, p. 28). A methodological path she will also follow later in her own book.

A chronological examination of these pioneers reveals a fascinating, yet flawed, attempt to grasp the Swahili metrical system using the strict accentuation patterns derived from classical Greek feet or other Western accentual traditions.

2.2.1. The Early Pioneers (1870–1920)

The initial foundational phase of this paradigm was built by a generation of pioneering philologists and missionaries who first encountered the poetry:

Edward Steere (1870): The earliest documented projection occurred in Steere's *Swahili Tales*. Confronted with an autonomous African poetic structure, Steere attempted to map it using classical

Greek prosodic models, specifically defining the *utendi* prosody as "**anapaestic**" (Steere, 1870, p. xi).

Carl Gotthilf Büttner (1894): Nearly twenty-five years later, Büttner continued the application of Greek models but arrived at a conflicting conclusion. He classified the Swahili *utendi* as fluctuating between "**iambic**" and "**trochaic**" meters (Büttner, 1894, p. XV).

Charles Sacleux (1909): Operating at the turn of the century, the French philologist, missionary of the Catholic church, and member of the Congregation of the Holy Spirit, Father Charles Sacleux shifted slightly away from strict Greek terminology, proposing instead a rhythmic model based on numerical beats and the syllabic count. He introduced the concept of the "oratorical accent", whose placement is dynamically variable: rather than merely reinforcing the regular tonic accent, it supersedes and replaces it in "emotionally charged speech" (Sacleux, 1909, p. 8), shifting away from the predictable penultimate slot to serve expressive needs. Sacleux assigned a specific, fixed accent count to each verse length: two accents for verses of three, four, six, and eight syllables; three accents for decasyllables; and four accents for verses of twelve and fourteen syllables (Sacleux, 1909, p. 327).

Conveniently, this fluidity of the oratorical accent acts as a methodological escape hatch: by allowing stress positions to shift dynamically under expressive conditions, it effectively shields the model from strict empirical verification and permits a form of selective sampling, immunizing by-design the theoretical framework against falsification.

The definition Sacleux gives of this accent deserves closer attention than it has received, because it introduces, in 1909, a distinction on which the present study turns. He states explicitly that the accent governing versification is not the tonic accent proper to each word taken in isolation, but the oratorical accent, which ordinarily assigns pitch and duration to the penultimate syllable of the principal words (Sacleux, 1909, p. 327). Three features of this definition matter.

The accent is not lexical by definition: Sacleux sets it apart from the tonic accent in the very sentence that founds versification upon it. It is selective, since it falls on the principal words, which means that it realizes a subset of the lexical stresses and not all of them. And it is free in placement, though not arbitrary: in his phonological introduction Sacleux explains that the oratorical accent either displaces the tonic accent and carries it elsewhere, or uses it and coincides with it on the same syllable, and that it may realize itself as pitch, as duration, or as intensity according to context, belonging to

emotionally charged delivery rather than to plain narration (Sacleux, 1909, p. 8).

The ordinary case is therefore coincidence with the penultimate, which is why Sacleux's accent counts look like counts of lexical stresses. But an accent that may depart from the lexical position, that falls only on the words the speaker treats as principal, and that belongs to delivery rather than to the text, cannot function as a generative law of composition. Werner, and the tradition that followed her, inherited Sacleux's figure of two beats without inheriting the definition that accompanied it, and treated as a metrical constant what its author had described as a feature of oratory.

Alice Werner (1917-20): Werner crystallized the core dogma of the accentual school. Her theoretical contribution largely consisted of taking these early observations and weaponizing them to bluntly dismiss the endogenous syllable count altogether. She stated: "[...] and the conclusion I am led to adopt is that the verse is measured by beats rather than by syllables" (Werner, 1917-20, p. 116, as cited in Vierke, 2011, p. 29). Like Sacleux, she posited a structure containing "exactly two beats" in every eight-syllable line, though she did so without his rigorous structural categorization of varying verse lengths (Werner, 1917-20, p. 116, as cited in Vierke, 2011, p. 29).

2.2.2. The Continuators: Consolidating the Accentual Tradition (1970s)

By the 1970s, the exploratory hypotheses of the early pioneers had naturally become the established academic baseline. Rather than questioning the foundational mechanics of the verse, a second generation of scholars—acting as dedicated continuators of this lineage—sought to refine and systematize these early observations.

J.W.T. Allen (1971): A towering figure in Swahili philology, Allen perpetuated the Greek analogy. In his metrical classification, he categorized the *utendi* prosody as "**ionic**" (Allen, 1971, p. 25, as cited in Vierke, 2011, p. 29).

A.M. Jones (1976): Approaching the poetry from a musicological perspective, Jones proposed a rigid accentual framework. He argued for a four-beat line consisting of regular "**trochaic feet**", where the major beats fall on the third and seventh syllables, and minor beats fall on the first and fifth syllables (Jones, 1976, p. 114, as cited in Vierke, 2011, p. 29, n. 35).

As Vierke correctly observes, this metrical classification was far from unequivocal, yielding highly conflicting definitions. The disagreement is usually presented as a matter of terminology — anapaest, iamb, trochee, ionic — and it is tempting to read it as the growing pains of a young field. It runs deeper than that. Jones hears four beats in the eight-syllable line, hierarchically ordered, with the strong ones on the third and seventh syllables and weaker ones on the first and fifth. Werner, and Vierke after her, hear two, on the third and the seventh. These are not two descriptions of a single structure; they are two different structures, and no procedure exists within the accentual tradition for deciding between them. A school that has been examining the same eight-syllable line for a century and a half, and cannot establish whether it carries two beats or four, has a difficulty that no refinement of terminology will resolve. Both positions are given a precise formal expression in section 2.5, where they can be stated in a way that allows them to be tested.

Jones's graded hierarchy deserves a further remark, because it is not an isolated eccentricity. Writing in the same period, Lyndon Harries records that Swahili verse contains "major (usually penultimate) and minor stresses" while denying, in the same sentence, that they constitute a metrical system of stress (Harries, 1962, p. 10). Two observers working independently, one a musicologist and the other an Arabist, describe graded prominence in the Swahili line. Only the two-beat model requires that prominence to be flat, and the graded reality it flattens will be measured directly in Chapter 9.

2.2.3. Metrical Excursus: The Anatomy of Exogenous Feet

When Western scholars invoke terms like "trochaic" or "anapaestic" they are deploying a vocabulary originally engineered for a completely different linguistic reality: the quantitative metrics of Ancient Greek. In classical antiquity, a metrical "foot" (πούς, *pous*) was not measured by the intensity or loudness of the voice, but by vocalic duration—the actual time it took to pronounce a syllable (long vs. short).

However, as poetry evolved within the English linguistic sphere—a language that do not possess a phonological opposition of vocalic length suitable for quantitative metrics—as well as the German linguistic sphere, which does feature such length distinctions, this Greek vocabulary was repurposed. European theorists emptied the

Greek terms of their quantitative substance and replaced it with an accentual-syllabic logic based on dynamic stress (the intensity of the voice).

Throughout this book, when the accentual school is said to use Greek feet, or a "classical Greek" trochee, it is these accentual redefinitions, as practiced in the Germanic languages, that are meant: the Greek name, applied to stressed and unstressed syllables rather than to long and short ones.

In the following pages, a hyphen (-) designates a stressed syllable, while a lowercase u designates an unstressed, weak, or neutral syllable.

In modern Western metrics, the Greek feet were redefined as follows:

The Trochee (Trochaic): From the Greek τροχαῖος (trokhaîos), derived from τρέχω (trékhō, "to run"). It originally meant a long syllable followed by a short one. In Western accentual metrics, it designates a stressed syllable followed by an unstressed syllable (- u). It produces a driving, marching, or running cadence.

The Iamb (Iambic): An unstressed syllable followed by a stressed syllable (u -). It mimics a natural walking pace and is the foundation of the English Shakespearean pentameter.

The Anapaest (Anapaestic): Two unstressed syllables followed by a stressed syllable (u u -). It creates a galloping rhythm.

The Ionic: A classical four-syllable foot consisting either of two stressed syllables followed by two unstressed (- - u u), or two unstressed followed by two stressed (u u - -).

In stress-timed languages like English or German, these meters function as absolute generative laws. The poet must actively engineer their syntax to ensure the natural lexical stresses of the words align perfectly with the predetermined percussion of the metrical grid.

To illustrate, consider the absolute perfection of the trochaic tetrameter (four trochees per line) in Henry Wadsworth Longfellow's *The Song of Hiawatha*:

> By the | shores of | Git-che | Gu-mee, By the | shi-ning | Big-Sea- | Wa-ter,

Here, the mechanical - u | - u | - u | - u pulse is relentless. The poet chooses every word specifically to feed this rhythmic engine. If a word's natural stress does not fit the trochaic beat, it is immediately discarded.

This accentual engineering is equally prominent in German, which further solidifies the rhythm through the addition of end-rhymes—a structural feature notably absent in Longfellow's unrhymed *Hiawatha*, as well as in classical Greek poetry. Consider the opening stanza of Friedrich Schiller's *An die Freude* (Ode to Joy), built on the exact same trochaic tetrameter:

Freu-de, | **schö**-ner | **Göt**-ter- | **fun**-ken,	(- u | - u | - u | - u)	[A]
Toch-ter | **aus** E- | **ly**-si- | **um**,	(- u | - u | - u | -)	[B]
Wir be- | **tre**-ten | **feu**-er- | **trun**-ken,	(- u | - u | - u | - u)	[A]
Himm-li- | **sche**, dein | **Hei**-lig- | **tum**!	(- u | - u | - u | -)	[B]

Here, not only does the natural dynamic stress of the German words perfectly coincide with the strong beats of the trochees, but the entire mechanical structure is tightly bound by an alternating A-B-A-B rhyme scheme (*funken/trunken*, *Elysium/Heiligtum*). This combination of strict accentual feet and end-rhymes represents the ultimate departure from Greek quantitative metrics, defining the standard machinery of classical German versification.

When this modern Western machinery is projected onto Swahili, a fascinating auditory illusion occurs, born from a collision between Swahili phonology and Western acoustic conditioning. Swahili possesses lexical stress, which is rigidly anchored to the penultimate syllable of the phonological word. Because every two-syllable word inherently carries stress on its first syllable, a poet will frequently—though coincidentally—construct an eight-syllable *utenzi* line using exactly four bisyllabic words.

Example: *Mtu huyu ana mambo* (This person has issues)

8 Syllables: **M**-tu | **hu**-yu | **a**-na | **mam**-bo

The acoustic illusion of a trochaic tetrameter: - u | - u | - u | - u

When scholars conditioned by centuries of Western accentual poetry encountered lines constructed in this specific manner, their ears eventually heard Schiller's *An die Freude* or Longfellow's *Hiawatha*. Convinced they had discovered the rhythmic core of the poem, they could declare the trochee the hidden generative law of Swahili verse.

This, however, is a profound methodological illusion: it elevates a property of the language into a universal metrical law. In a true trochaic system, the rhythmic pulse commands the vocabulary. In Swahili, the exact opposite is true: the vocabulary commands the

accents, while only the total syllable count (*mizani*) remains sovereign. The absolute absurdity of applying the trochaic label to Swahili becomes spectacularly evident the moment the poet selects words of different lengths. If a line contains a five-syllable word followed by a three-syllable word (e.g., *a-li-si-**ki**-a sa-**u**-ti* [(s)he heard a voice]), the natural penultimate stress falls on the 4th and 7th syllables u u u - u | u - u. The beautiful, marching - u | - u | - u | - u percussion instantly shatters, unless the linguist or philologist saves the phenomena with an explanation that maintains the trochaic rhythm where none exists.

As will be explored in detail later in this work, contemporary scholars of the accentual school are not the architects of this auditory mirage, but rather the heirs to a century-old theoretical tradition founded upon it. In managing this inherited framework, their modern impositions onto Swahili poetry diverge along two distinct paths.

On one hand, Clarissa Vierke (2011) proposes a framework she classifies as *gemischt-quantifizierend* (mixed-quantifying). Her model is "mixed" precisely because it combines strict syllabic counting and syllabic end-rhymes with a system of stress patterns. Operating within the concept of the *Grenzschema*, these stress patterns are not entirely trochaic, only their nuclei (see Section 2.3.1); rather, they rely on two main beats per line, echoing the earlier work of Alice Werner. This specific structural synthesis owes a great deal to the mechanics of classical German metrics.

On the other hand, Minerba (2025) represents a more conservative return to Greek metrical terminology and the legacy of early scholars, though he arrives there through a completely different paradigm: diffusionism rather than straightforward Eurocentrism. Minerba initially attempts to project an Arabic metrical framework onto the Swahili verse. When this diffusionist projection ultimately fails to map onto the empirical linguistic reality, he falls back on Hellenistic terminology, conveniently categorizing the rhythm as a "trochaic innovation."

Ultimately, whether relying on a German-inspired mixed model or defaulting to Greek trochaic slots after a failed Arabic projection, both approaches force the Swahili verses into a Procrustean bed. Driven by a variant of the Exo-African Postulate, they elevate their preconceived metric frameworks above empirical observation, systematically contorting a distinct linguistic reality to fit their models.

2.3. Clarissa Vierke and the Application of the *Grenzschema*

To properly evaluate the accentual hypothesis, it is intellectually imperative to examine its most sophisticated contemporary articulation. In her 2011 critical edition and stylistic analysis of the 19th-century *Utendi wa Haudaji*, Clarissa Vierke undertakes a massive and largely admirable philological endeavor. Before deconstructing her metrical conclusions, one must acknowledge the care with which she defines her corpus and its limitations. Vierke explicitly recognizes the profound historical rupture in Swahili manuscript culture, defining her approach as a form of "text archaeology". She concedes from the outset the extreme methodological difficulty of her task: there are no surviving historical reports detailing how *tendi* were composed, performed, or transcribed in the 19th-century, nor does any prescriptive endogenous poetics from that era exist.

Furthermore, she is fully transparent about the fact that she does not rely on a single manuscript. Her edition is a rigorous reconstruction based not only on the transliteration executed by Ernst Dammann and Muhamadi Kijuma in the 1930s (manuscript DA), but also on two additional manuscripts in Arabic script from the J.W.T. Allen collection (A1 and A2). Lacking access to the original 19th-century performative context—an inevitable constraint as recordings were not yet extant—she relies heavily on the textual data and the modern "reading and recitation" and "chanted renderings" by contemporary scholars—specifically Ahmed Sheikh Nabhany, Khadija (Mama Dida), and Ahmad Nassir—to deduce the historical prosody of the text (Vierke, 2011, pp. 23-24).

However, in attempting to unearth the generative engine of the classical *utendi*, Vierke makes a decisive and highly problematic theoretical pivot. It is here that we witness the precise mechanics of the epistemological marginalization of the endogenous Swahili voice. She dutifully records her Swahili informants' uncompromising insistence that the meter is governed exclusively by an absolute

syllable count, but she immediately sidelines this consensus in favor of an exogenous rhythm. She writes:

> Number of syllables (*mizani*) and rhyme (*vina*) are two parameters which have frequently been pointed out in the literature on *tendi* as the most essential prosodic characteristics. Nabahany and Ahmad Nassir would also underline the two parameters as quintessential. From their perspective, the line meter is primarily based on syllable count: any deviation from the prescribed eight-syllable count is a fault (*guni*) that needs correction and a sign of incompetence (Abdulaziz 1979: 60). While rhyme and the quantity of eight syllables per line have been considered the prosodic hallmarks of the *utendi*, much less attention has been paid to rhythmicisation through accent or beat – and where this was discussed it raised more controversy. Ever since Steere (1870) in his 'Swahili Tales', scholars, especially at the beginning of the 20th century, have tried to grasp the prosodic system in terms of strict accentuation patterns derived from their experience with classical Greek texts and their supposed strict metrics. (Vierke, 2011, p. 28)

This passage encapsulates the foundational reflex of the accentual school. The Swahili master poets state that their art is governed by *vina* and *mizani*, and that any deviation from the eight-syllable count is a severe structural failure (*guni*). Yet, instead of accepting this explicitly articulated, formally perfect endogenous and autonomous theory as the definitive generative law of the verse, Vierke brushes it aside. She aligns herself directly with the exogenous lineage of Edward Steere, Alice Werner, and Charles Sacleux, resurrecting the pursuit of "rhythmicisation through accent or beat".

2.3.1. Visualizing the Accentual Edifice: Key Concepts of Clarissa Vierke's Metrical Theory (2011)

To fully grasp the mechanical complexity of Vierke's framework before deconstructing it, it is helpful to visualize its core components. Firmly situated within the tradition of the accentual school, her theory builds directly upon Alice Werner's bipartite, two-beat model. By leveraging the natural penultimate lexical stress of the Swahili language, Vierke isolates specific trochaic nuclei [- u] (a stressed

syllable followed by an unstressed one). However, because this trochaic nucleus strictly accounts for only two syllables, she is forced to categorize any preceding, unaccented syllables as a highly variable and metrically irrelevant "pre-nucleus" Finally, to reconcile this volatile accentual rhythm with the undeniable reality of the traditional eight-syllable line (*mizani*), she proposes a hybrid, mixed-quantifying mechanism (*gemischt-quantifizierendes System*). This hybrid framework must also account for the absolute structural requirement of *vina* (rhyme).

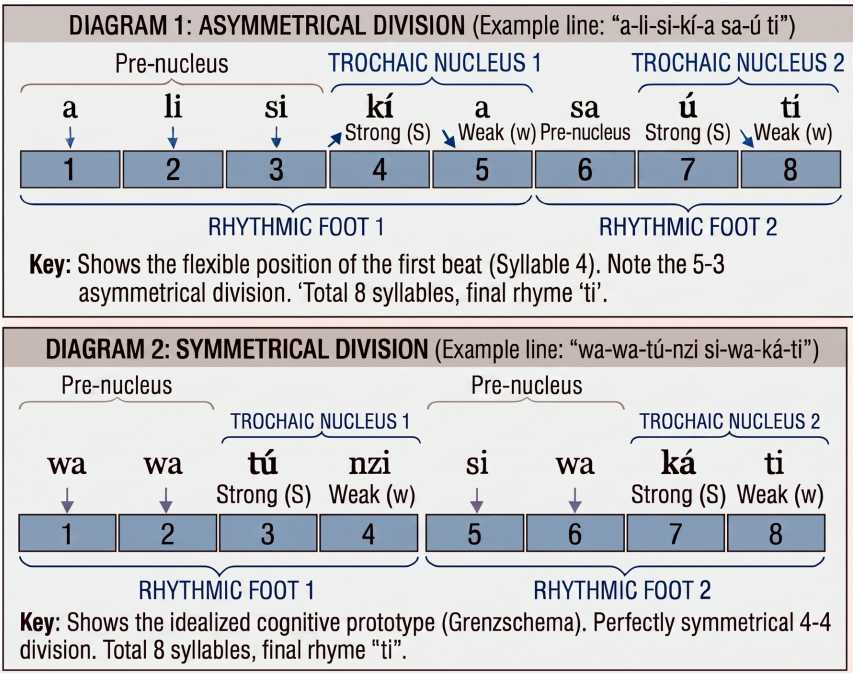


Figure 1: *Clarissa Vierke's metrical theory of the utenzi meter (2011): extended presentation and key concepts*

The diagram illustrates this integration using two *utenzi* 8-syllable lines that share a final rhyme in *-ti*: *alisikia sauti* [He/she heard a voice] and *wa watunzi si wakati* [It is not the composers' hour][2]. At the apex of this theoretical system sits the *Grenzschema*—an

[2] These illustrative lines were composed for this study specifically to demonstrate this metrical mechanism.

idealized, perfectly symmetrical 4-4 cognitive prototype designed to project a sense of structural order onto the underlying syllabic asymmetry.

1. Key Metrical Concepts

1.a. Bipartite Two-Beat Accentual Model (Based on Alice Werner)

Line structure is defined by exactly two primary beats per line, overlaid in fixed 8-syllable grid. A fixed terminal beat anchors the penultimate syllable (Diagram 1, S7). A second flexible beat is in the first half of the line (Diagram 1, S4).

1.b. Trochaic Nucleus Structure

Each beat forms, with the syllable that follows it, a trochaic nucleus [-u], consisting of a stressed (Strong) syllable and an unstressed (Weak) syllable. This structure leverages natural Swahili penultimate lexical stress.

1.c. Pre-Nuclei

Variable syllables located outside the core trochaic nuclei. Considered metrically irrelevant in Vierke's theory, these elements provide flexibility without altering the fundamental two-beat rhythm.

1.d. The *Grenzschema*

An ideal 4-4 symmetrical pattern, the *Grenzschema*, serves as a cognitive prototype of the ideal prosodic pattern.

1.e. Mixed Metrics (*gemischt-quantifizierend*)

The accentual metrics is superimposed onto the classical Swahili metrical system, which relies strictly on syllable counting (*mizani*) and rhyme schemes (*vina*).

2.3.2. The Mechanics of the "Swahili Foot" and the Trochaic Hypothesis

Having bypassed the absolute sovereignty of the syllable count, Vierke sets out to construct her accentual model. It is an undisputed phonological fact of the Swahili language that lexical stress naturally falls on the penultimate syllable of a word. Because classical Swahili

poetics strictly prohibit enjambment, the end of the line coincides with the end of a lexical and syntactic unit, and the final word of the line therefore carries its own lexical stress on its penultimate syllable. In an eight-syllable line whose final word has two syllables or more — which is to say in every regular line of the corpus — that stress falls on the seventh position. The invariant is thus lexical before it is positional: the terminal stress is the penultimate of the last word, and it lands on syllable 7 as a consequence of the syllabic constant rather than as a metrical requirement. Stated in this form it holds, by the lexical rule of the penultimate stress[3], without a single exception, across the corpora analyzed in Chapter 9, including in the *guni* lines, where it falls on syllable 6 or 8 according to the length of the line. Vierke identifies this salient seventh syllable as the primary, fixed beat of the line.

Building upon this fixed terminal anchor, she postulates the existence of a second, less prominent beat occurring in the first half of the line. She argues that the eight-syllable line is inherently bipartite, consisting of exactly two syntactic-prosodic units, which she terms "rhythmic groups", "feet" or "colons" (Vierke, 2011, pp. 31-34). Because there are exactly two accents per line, the line must logically be divided into two feet.

It is precisely at this juncture that the accentual theory encounters a severe structural paradox. While the second beat is permanently locked onto the seventh syllable, the position of the first beat is highly flexible, dictated entirely by the natural syntax of the phrase. It may fall on the first, second, third, or fourth syllable. Consequently, the two "feet" that constitute the line are rarely symmetrical. The actual texts routinely produce highly asymmetrical divisions: a 2-syllable foot followed by a 6-syllable foot (2-6), a 3-5 division, a 5-3 division, or even a 6-2 division (Vierke, 2011, pp. 31-34).

To clarify a frequent point of terminological confusion—specifically regarding the notion of a "trochaic" meter—one must understand how Western metricians classify these varying foot lengths. As seen before (see 2.2.3), in classical Western metrics, it is essential to distinguish between the metrical foot (the basic rhythmic unit) and the meter (the

[3] Our testing protocol operates as a three-step funnel: by rule from the text, by ear from a formal reading, and by instrumental measurement. As we shall see, instrumental checks on sample recordings fail to fully corroborate the supposed stability of this penultimate position. Stress can instead fall on segment boundaries—landing, for example, on the eighth syllable of an eight-syllable hemistich (see Chapter 9 for a detailed analysis).

overarching structure of the line). A "trochee" is a specific metrical foot consisting of exactly two syllables: one stressed syllable followed by one unstressed syllable [– u]. Because Swahili possesses a natural penultimate stress rule, any isolated two-syllable word automatically forms a trochee [– u]. These individual feet can then be combined to establish a complete meter, such as a tetrameter, which repeats this pattern four times [- u] · [- u] · [- u] · [- u]

When Vierke analyzes a highly asymmetrical line division, such as the 2-6 split, the first foot is composed of only two syllables, meaning its stress must fall on the first of the two [– u]. Therefore, the first foot is technically a trochee. In her footnotes, Vierke explicitly addresses this, noting that J.W.T. Allen rejected the 2-6 division precisely because he believed the Swahili meter was never trochaic:

> Allen (1971: 28) would rule out the 2-6 division, as he points out that the meter is never trochaic, whereas the first foot in a 2-6 division would necessarily be trochaic" (Vierke, 2011, p. 31, n. 42).

Vierke, countering Allen, argues that the 2-6 division (and thus, an initial trochaic foot) is indeed possible in performative practice.

Furthermore, regarding the exact placement of this first accent, Vierke makes claims that seemingly contradict standard Swahili phonology, where isolated monosyllabic words do not carry primary stress. She argues, without offering a means of verification, that the first beat can indeed fall on the very first syllable of the line. To justify this, she relies on the merging of enclitics and proclitics into larger phonological phrases:

> In the first half, besides falling on the third syllable, accent might also fall on the second (kulá mi liyuwikane (DA 379) or on the first syllable (kwá ye Abalhasani (DA 843) and sometimes even on the fourth syllable (nakuombá we Rabana (DA 804), so that feet tend to vary than remaining stable in the performative practice of *tendi*. (Vierke, 2011, p. 31)

She explicitly rules out a 7-1 or 1-7 division, however, because the penultimate stress rule requires that every foot must contain at least one unaccented syllable following its accent-bearing syllable.

2.3.3. Theoretical Adjustments: Arabic Cores and "Irrelevant" Syllables

To a traditional metrician, a system where the constituent "feet" randomly fluctuate between two and six syllables is the very definition of arhythmia; it mathematically precludes the existence of a regular, repeating metrical beat. To save the accentual phenomenon and reconcile this massive syllabic irregularity with the concept of a structured rhythm, Vierke introduces a highly elaborate epistemological epicycle borrowed directly from the study of pre-classical Arabic prosody, explicitly confirming the Exo-African Postulate:

> This analysis draws much on Weil's analysis of Khaliil's system of pre-classical Arabic versification schemes, where stress-accent played an important role (Weil, 1958). Like in Swahili *utendi*, stress accent falls on a long syllable which together with an adjacent short syllable constitutes the rhythmic nucleus (Weil, 1958: 36). Weil (1958: 36; 60) describes the Arabic foot as consisting of two parts: on the one hand, a rhythmic core (*autaad*) which invariably bears the accent and which is quantitatively fixed in the sense that an unchangeably long syllable is preceded by an unchangeably short syllable and on the other hand, a rather neutral periphery (*asbaab*) which is unspecified with respect to syllable number and mora count, but never bears accent." (Vierke, 2011, pp. 33)

By forcing the Swahili syllabic grid into this Arabic logic, Vierke dissects the Swahili foot into a bi-syllabic "rhythmic core" and a periphery of "weak" or "neutral" syllables whose absolute number is deemed metrically irrelevant. She pushes this logic to its absolute extreme, declaring a total rupture with the Swahili concept of *mizani*:

> Thus, building on Jakobson, we can say that in contrast to systems of versification based on a strict musical rhythm and count of syllables, in 'accentuating metrics' the sheer number of syllables is irrelevant. With respect to the *utendi*, line-internally it is not the number of syllables of each foot which is of relevance, but rather the two accents which oppose each other and which account for the rhythm. (Vierke, 2011, p. 32)

And further:

> Thus, the number of syllables is irrelevant as a principle, and it is the accent and the invariable, unchangeable rhythmic nucleus which – like a rock – constitutes the rhythm of the line." (Vierke, 2011, p. 32)

This is a devastating theoretical leap. To state that "the number of syllables is irrelevant as a principle" in a poetic tradition whose practitioners explicitly define their art through *mizani* (measure/syllable count) is an extraordinary invalidation of the endogenous intellect. The syllables are stripped of their arithmetic equality (*urari wa mizani*) simply to ensure that the highly asymmetrical textual data can be forcefully accommodated by an imported rhythmic theory.

However, the sheer physical reality of the manuscript—the unyielding fact that every single line still rigidly conforms to exactly eight syllables—forces Vierke to immediately backtrack and introduce yet another theoretical patch. Recognizing that the Arabic model of boundless "weak" syllables cannot fully explain the absolute eight-syllable limit of the *utendi*, she creates a hybrid categorization:

> However, the variability in syllable number is a feature that holds only line-internally: the *utendi* line invariably consists of eight syllables providing a rigid frame to variability. In that sense, following Weil (1958), we can consider the *utendi* as making use of a mixed-quantifying metrical system (*gemischt-quantifzierend*), which employs both stress-accent working around a rhythmic core and the syllable quantum. (Vierke, 2011, p. 34)

This "mixed" system (*gemischt-quantifizierend*) is the ultimate epicycle. The theory imposes an accentual framework, dismantles the relevance of the syllable count to make the accents fit, and then retroactively re-imposes the syllable count because the empirical data (the 8-syllable line) refuses to disappear.

2.3.4. The *Grenzschema*: Evaluating the Abstract Cognitive Prototype

Having constructed this complex mechanism, Vierke then proceeds to define the ultimate aesthetic ideal of her accentual system: the *Grenzschema*.

She theorizes that despite all the variations, the third syllable frequently operates as the favored position for the first beat. When the stresses align exactly on the third and seventh syllables, the eight-syllable line perfectly bifurcates into two symmetrical, four-syllable feet: [u u - u] · [u u - u]

In discussing the structure of the *tendi* meter, Vierke introduces an observation based on the oral testimony of practitioners. She states:

> When I asked Ahmad Nassir and Ahmad Sheikh Nabahany about well-formed lines in *tendi*, I found them highlighting the lines which display a symmetric structure of two four-syllable feet, which according to them, displayed an ideal prosodic pattern. (Vierke, 2011, p. 34)

This reliance on hearsay presents a significant methodological flaw, compounded by the complete absence of the original transcript. The reader is left without the original Swahili phrasing, the exact terminology used by the poets, or even the knowledge of whether the conversation occurred in Swahili or English. Consequently, the concept of "an ideal prosodic pattern" is entirely mediated by the author's translation and interpretation, severed from its original linguistic context. Moreover, even if one takes this transcribed phrasing literally, the poets explicitly refer to "an ideal prosodic pattern." This merely acknowledges a preferred aesthetic among others, rather than establishing this symmetry as the exclusive, binding rule of the meter.

Furthermore, when attempting to illustrate this theoretical symmetry, the methodology relies heavily on selective sampling without offering a means of empirical verification. The author continues:

> Another hint at the ideal prosodic pattern comes from regularly recurring formulas. We can certainly assume that symmetry in structure, in semantics as well as prosody, make a string of words, a formula, most easily memorisable. The clustering of features

> guarantees a formula's stability. In the following stanza, DA 301, the formula *kwa zivúmi na siáha* 'in bluster and tumult' occurring in the second line ideally represents the bipartite division of the line into two equal feet. Stress falls on the third and seventh syllables. (Vierke, 2011, p. 34)

By isolating a single line— out of a corpus comprising hundreds of lines where stresses are left unmarked —that conveniently corroborates the hypothesis, the demonstration falls short of methodological rigor. The text fails to specify exactly how these stresses were obtained. It remains entirely unclear whether the verse was analyzed as it is sung, chanted, or simply read aloud. While applying a normal spoken utterance does indeed place the stresses on the third and seventh syllables in this specific instance, the absence of an explicitly defined oral realization modality leaves open the possibility that reading styles are being conflated or shifted for theoretical convenience. Because no statistical analysis regarding the actual distribution of this pattern across the broader corpus is provided, it is impossible for the reader to verify the claim or reproduce the experiment.

The theoretical escalation in this argument is quite striking. What initially emerges as merely one subjective prosodic preference among others, and is subsequently justified by a single, methodologically opaque line cherry-picked from a corpus of hundreds, is ultimately transformed into an absolute psychological reality. Building upon this rather fragile empirical foundation, Vierke elevates this 4-4 symmetry—specifically the bipartite structure [u u - u] · [u u - u]—to the status of an abstract cognitive prototype, borrowing Roman Jakobson's concept of the *Grenzschema* (the boundary schema):

> Jakobson (1933), talking about the *desterac*, where he finds an ideal-typical though not necessarily realised pattern of five syllables per colon, calls it a *Grenzschema*, which manifests itself in every line and which the poet has in mind, 'even though he might not be able to abstract and to define it' (Jakobson 1933: 50). The *Grenzschema* in Jakobson's study is an abstraction in the sense that it leaves open possibilities for variation– comparable to the variability that we have found in the make-up of the feet of the *utendi*. (Vierke, 2011, p. 37)

This admission—that the *Grenzschema* is an "abstraction" and an "ideal prosodic scheme" that is "hardly ever completely realised"—exposes the deepest empirical void of the accentual school and its immunization against experimental testing against data. A scientific law of poetics must be observable and predictive. If the generative engine of the verse is a [u u - u] · [u u - u] accentual pulse, that pulse must manifest consistently. Yet, the theory explicitly relies on confirmation bias, isolating non-dominant patterns and elevating them to the status of a generative rule.

The fragility of this theoretical edifice becomes glaringly apparent immediately following the introduction of the *Grenzschema*, where the author inadvertently introduces a striking double contradiction. In a revealing passage, she states:

> However, as seen above, prosodic symmetry as well as accentuation following the syntax of normal speech can be made to coincide, as is most obvious in the hendiadiyonic formulas highlighted by Nabahany and Ahmad Nassir for their beauty, i.e. their perfect analogous symmetry in prosody, syntax and semantics. Both Ahmad Nassir and Nabahany were not in favour of an imposed strict musical rhythm that always falls on the third syllable of a foot, as it was against the *utendi*'s narrative style and–as we could add–its closeness to the natural flow of speech. (Vierke, 2011, p. 37)

This admission dismantles the argument established earlier. How, then, is the reader supposed to reconcile these statements? If, on page 34, these very same poets were invoked as foundational authorities to establish this symmetric stress pattern as the absolute "ideal prosodic pattern," how can they now be cited as actively rejecting the imposition of that exact rhythm?

First, the author clearly contradicts her own prior demonstration by acknowledging that the practitioners oppose the strict application of the model she just constructed. Second, she creates a profound methodological paradox by contradicting the practitioners themselves. By continuing to formalize this specific stress pattern into an overarching cognitive *Grenzschema*, the author overrides the explicit aesthetic principles of the Swahili poets. She relies on their oral testimony to legitimize the creation of her theoretical model, only to dismiss their explicit opposition when their reality no longer aligns with her framework.

This methodological disarray is further exacerbated by the author's failure to define the oral realization modality of her single example, DA 301. As previously noted, the text does not specify whether this verse is analyzed as a normal spoken utterance, chanted, or set to a musical melody. This is not merely an issue of selective sampling; it introduces a profound rhythmic chaos into the analysis. In Swahili, a normal spoken utterance is strictly governed by the deterministic rules of Swahili phonology and syntax. Conversely, a chanted or musical performance introduces highly variable, unpredictable rhythmic fluctuations dictated by the chosen melody.

To bring order to this epistemological confusion and establish a framework for reproducible experimental analysis, it is imperative to conceptualize the Swahili poem through a tripartite formal architecture. This structural distinction, which we define as the "3-Layer Theory of Swahili Poetics" allows for a rigorous classification of poetic phenomena by strictly isolating the types of oral realization (see 3.3 for a detailed explanation). At its absolute base is Layer 1 (The Syllabic Law), the generative foundation of the poem governed strictly by *vina* (the rhyme boundaries) and *mizani* (the syllable count). Upon this foundation rests Layer 2 (The Macroscopic Architecture), which organizes these syllabic segments into broader structural units such as lines and stanzas and specific classical forms. Finally, at the opposite end, is Layer 3 (The Melodic Overlay), the domain of the singer and the musical melody, which introduces the variable and chaotic rhythmic fluctuations of oral performance.

Crucially, lexical stress operates outside this metrical engine: it is a property of the Swahili language, not a property of the meter. While it inevitably surfaces in the acoustic reality of vocalization, it holds no generative function. By failing to isolate these levels, the entire theoretical edifice of the accentual school collapses into a performative fallacy: they have consistently confused a phonetic property of the spoken language—magnified within the variable acoustic chaos of Layer 3—with the metrical rules of Layers 1 and 2.

Another recurring tendency in Vierke's work (2011) is the scarcity of actual empirical proof, which is systematically compounded by a circular reliance on evidence that selectively confirms her own theoretical model. Across hundreds of pages of her extensive analysis, Vierke provides explicit accentual markings for only a very limited selection of highly specific verses. But when she needs it, all of them are strictly conforming to the "two beats" by line rule she explicitly borrowed to Alice Werner.

The single most prominent example she offers to demonstrate this binary rhythm is:

> To exemplify this variation, we can take a stanza from the Haudaji, as for instance DA 130, where I have marked the occuring beats: *Kitúzo cha mato yángu / póza la nafusi yángu / rasúli, thumwa ya Múngu / mwényе kutushufaQía?* (Vierke, 2011, p. 29)

Once again, we are given no indication that would allow us to reproduce her conclusion. In her marking, a line such as "*Kitúzo cha mato yángu*" carries two beats, dividing it 3 + 5. This fits her model: the 4 + 4 *Grenzschema* is an ideal that is not always realized, but the two-beat rule itself must be realized in every line of the *utenzi* meter as Vierke conceives it. The lexicon, however, assigns this line three stresses, since the deterministic lexical rule places stress on the penultimate syllable of every phonological word, monosyllabic grammatical words being proclitics: *Kitúzo cha máto yángu*. A reading may weaken one of these three stresses, but only a reading could show it.

This raises methodological questions. Where does the data originate? How was the line performed, read or sung? And why does her marking give no beat to the noun *mato*, the Kiamu variant of standard *macho* ("eyes"), while giving one to its determiner *yángu* ("my")? Without the reading behind the marking, the reader has no means of checking it, and must take the author's word for it.

To base an entire metrical paradigm on a handful of carefully selected, marked examples while leaving the vast majority of the corpus unmarked is to substitute a statistical artifact for a formal law. The ultimate collapse of this theoretical edifice is quietly confessed by Vierke herself in a devastating footnote regarding the realization of this accentual rhythm in actual performance:

> I have no evidence for an *utendi* where the reciter completely adhered to a strict rhythmic pattern –which again emphasizes its status as a prototype rather than a canonic norm. From the meagre evidence that we have, we can at least presume that variation in accentuation existed in the 19th century [...] (Vierke, 2011, p. 38, n. 47)

If there is "no evidence" of a reciter adhering to this strict rhythmic pattern, and if the pattern is "hardly ever completely realised" in the text itself, what status does it hold? Can a generative metrical law operate entirely detached from actual textual and oral realization? One might venture to suggest that, as a generative metrical law, the pattern may simply not exist. It is probably a phantom.

2.3.5. The Methodological Shift Toward Performance and the Musical Rationale

Finally, to justify how this purported accentual rhythm is sustained, Vierke leans heavily into a performative fallacy. She attempts to isolate a specific performative mode by drawing a distinction between *kuimba* (singing, which she admits imposes its own rhythm and overrides speech intonation) and *kukangaya* (chanting or scanning). As we have already observed in the introduction, the word *kukangaya* is absent from the dictionaries, and the precise mediality of its oral realization is unclear to Vierke, as she explicitly acknowledges. What is important here is that this mediality operates strictly at the level of performance, as a chanted mode.

To legitimize this performative shift, Vierke explicitly relegates the normal spoken utterance to a subordinate status:

> Besides the chanted mode, there is also a reading or spoken mode of *utendi* actualisation [...] The reading mode has to be considered as a secondary mode of recitation, as *tendi* are composed to be chanted as part of a performance and not to be read out. (Vierke, 2011, p. 24)

However, this justification rests on profound lexicographical and cultural contradictions and a conceptual error. On the one hand, Vierke commits a profound ontological conflation regarding the nature of the text itself. It is historically accurate that *tendi* are composed to be sung; yet, their primary material existence is graphic. As will be demonstrated, this written dimension introduces a graphic reason—a spatial logic that organizes rhyming chains along a strict syllabic grid. By illicitly inferring that the written manuscript is merely an ancillary script for oral delivery, Vierke treats the teleological outcome (the song) as the structural genesis of the meter. This ontological inversion of text and performance is a logical error already present in Steere (1870) and uncritically perpetuated here.

On the other hand, Vierke openly admits that she "could not find *kukangaya* [...] in any of the Swahili dictionaries" (Vierke, 2011, p. 24, n. 1). Yet, in the very same footnote, she summarily dismisses the term *kughani*, erroneously claiming it refers exclusively to Arabic Quranic recitation.

In reality, as demonstrated extensively in the introduction, this claim is refuted by no fewer than four standard reference dictionaries (including Wamitila, TUKI, BAKITA, and Sacleux). These authoritative sources universally codify *kughani* (and its agent noun *mghani*) as the endogenous, accepted Swahili term utilized by the practitioners themselves for singing or reciting poetry, proving her dismissal of the word to be entirely unfounded.

By contesting the standard, universally documented endogenous terminology in favor of an obscure, undocumented category, Vierke artificially carves out a highly specific, yet entirely nebulous performative space. Epistemologically, this maneuver serves only to immunize her accentual theory against falsification. It becomes apparent that the methodology lacks the conceptual and empirical tools necessary to conduct a rigorous, reproducible experiment.

From a strictly Popperian perspective, a scientific metrical law must be based on stable, observable, and falsifiable forms. Just as an eight-syllable line consistently manifests exactly eight syllables in every printed line, a generative accentual law must demonstrate predictable regularity. If her theory of exactly two accentual feet per line or the *Grenzschema* were a genuine metrical law, rigorous methodology would dictate, according to Vierke's own standard, having a performer chant the epic's hundreds of lines in their entirety, followed by a systematic statistical mapping of every accent to prove a regular distribution. She does no such thing.

Because the chanted oral realization—anchored in oral performance—inevitably crystallizes various, non-deterministic accentuation schemes depending on the chosen melody, she cannot provide a unified dataset. Instead, we are left with no means of verification, as she provides accentual markings for only an isolated line or two that conveniently happen to align with her preconceived two-beat model. It is precisely because the empirical data fails to systematically support her claim that she is forced to rely on the abstract concept of the *Grenzschema*.

By cherry-picking these rare performative instances, she completely bypasses the deterministic, syllabic reality of the written grid. As will be formally established later (see section 3.3), this

methodological error consists of anchoring the analysis in the non-deterministic acoustic fluctuations of oral performance (Layer 3), rather than grounding it in the invariant architecture of the text itself (Layers 1 and 2). Consequently, the structural blueprint is verified on the text, and the accentual claims are tested on normal, stable readings, which strip away the melodic distortions of performance.

Ultimately, the accentual school attempts to save the appearance of a Western accentual beat at the direct expense of the African intellect. The result is a highly sophisticated metrical epicycle designed to explain away the observable reality of the Swahili manuscripts and their standard oral realization. In this theoretical contortion, Swahili-language poets and scholar are instrumentalized as epistemic alibis: they are summoned to the academic stage to legitimize the model when their vocabulary aligns with the hypothesis, only to be abruptly silenced and dismissed the moment their explicit testimony contradicts the researcher's preconceived framework.

2.4. Emiliano Minerba and the Comparative Approach: The Diffusionist Hypothesis

The culmination of the exogenous metrical theories finds its most contemporary and sophisticated articulation in the work of Emiliano Minerba (2025). While the earlier generations of the accentual school—represented by scholars from Alice Werner to Clarissa Vierke—relied on the direct projection of classical Greek or Germanic rhythmic paradigms onto Swahili verse, Minerba subtly shifts the epistemological battleground. He attempts to salvage the accentual hypothesis by embedding it within a broader, highly erudite comparative framework, tracing the diffusion of Arabic prosody across the Islamic literary networks of the African continent. This diffusionist hypothesis posits a complex historical creolization between imported Islamic forms and two African languages—Wolof and Swahili—to explain the architecture of Wolof and Swahili classical meters.

However, as the following analysis will demonstrate, framing Swahili poetics within an Arabo-centric historical narrative does not resolve the fundamental contradictions of the accentual school. While Minerba correctly acknowledges that strict quantitative Arabic meters are structurally incompatible with Swahili verse, he paradoxically fills

this metrical void not by turning to the endogenous syllabic principles of the language, but by relying on a Western metrical framework. By systematically dissecting his methodology, we will observe that rather than breaking away from the Western accentual tradition, his diffusionist approach acts as its most faithful contemporary continuation. Firmly anchored in the academic lineage established by Alice Werner and Clarissa Vierke, Minerba uses the narrative of cultural hybridization simply to transpose and legitimize their foundational rhythmic premises. Through an examination of his fundamental linguistic premises (2.4.1), the dangers of his macro-extrapolations between Wolof and Swahili (2.4.2), the limits of his Ḥumaynī structural thesis (2.4.3), and finally, his ultimate revival of the phantom trochee (2.4.4), this section will expose how this imposing historical superstructure relies upon the exact same epistemological flaws as its predecessors: the manipulation of a passive phonological inevitability to engineer a Texas Sharpshooter fallacy, and the perpetuation of a relentless performative fallacy.

2.4.1. A Necessary Preliminary: Reading Minerba as He Actually Is

Academic rigor demands, before all else, that we confront the position our interlocutor actually occupies rather than a simplified strawman. The framing of Emiliano Minerba's *Meters in Motion: History of Swahili and Wolof Prosodies* (2025) as a blindly exogenous work, or as a wholesale, uncritical application of Arabic quantitative metrics to classical Swahili poetry, would be an inaccurate summary—and an intellectually dishonest target.

Minerba's work represents the most recent, sophisticated, and erudite iteration of the accentualist lineage. He approaches the Swahili and Wolof corpora with an impressive mastery of the Arabic *ʿilm al-ʿArūḍ* (the science of metrics), accurately describing its highly technical mechanisms—such as the *ḥarf* parsing, the *sabab* and *watid* elements, and the variations of *ziḥāfas* and *ʿillas* (Minerba, 2025, pp. 116-137). He sets out his comparative agenda clearly from the introduction:

> "In some domains, such as stanzaic structure and the use of rhyme, the similarity with the Arabic system has been outlined [...] On the other hand, it is understandably more difficult to find similarities in the internal structure of Swahili and Arabic metres:

> the first are based on isosyllabism and the placement of stresses throughout the line, whereas the second are quantitative (namely, based on syllabic length) and not constrained to a constant number of syllables." (Minerba, 2025, p. 2)

Regarding the phonological bedrock of the language, Minerba explicitly concedes a linguistic reality that serves as a fundamental prerequisite for this manifesto's overarching argument against exogenous models: the absolute impossibility of applying Arabic quantitative metrics to the Swahili language. While our primary target remains the fallacy of accentual metrics, Minerba's necessary admission regarding the absence of vocalic quantity closes the door on the direct importation of the Arabic quantitative model. He states this with clinical clarity:

> It is linguistically impossible to employ a quantitative metre such as the Ḫalīlian one in Swahili. Unlike Arabic, in fact, the Swahili language does not distinguish between vowel lengths, and its syllables cannot end with a final consonant. (Minerba, 2025, p. 151)

He further acknowledges that for both the archaic and the classical periods, "the fundamental element of the metrical pattern is the *syllable* [emphasis added]" (Minerba, 2025, p. 149), correctly noting that "'Isosyllabism' is expressed in Swahili as *urari wa mizani* ('equality of the measures')" (Minerba, 2025, p. 156)

On this critical phonological point, there is no disagreement between his empirical observations and our data-driven approach. However, this concession constitutes the narrow ground upon which a far more subtle, and therefore more dangerous, set of epistemological maneuvers is erected. To save the accentualist framework inherited from the Werner-Vierke lineage of the accentual school of Swahili classical poetry, Minerba deploys a "comparative approach" that reframes the development of Swahili poetics as a process of cultural hybridization and appropriation:

> Using the comparative approach, I will claim that both Swahili and Wolof literatures have developed as a mix of indigenous and Arabic elements, in a process of appropriation that saw the receiving communities—the Swahili and the Wolof—not as passive imitators, but as conscious managers of the influence they were exposed to. (Minerba, 2025, p. 6)

As the following analysis will demonstrate, while this framing sounds culturally respectful, it inadvertently forces classical Swahili poetry into an evolutionary paradigm that ultimately violates its endogenous formal reality. We will see how this approach strips the Swahili poet of their formal sovereignty by reintroducing the phantom of the accentual foot through the back door of historical derivation.

2.4.2. The Limits of Extrapolation: *Wolofal*, Classical Swahili Poetry, and the Risk of a Double Erasure under Arabic Projection

Minerba's project explores the history of Swahili and Wolof prosodies by invoking the Arabic *ʿArūḍ* to understand rhythmic patterns across a broader African Islamic continuum. However, the severe limits of this comparative approach are immediately exposed by a fundamental difference in the phonetic equipment of these two languages—a chasm that Minerba himself documents but whose epistemological consequences he bypasses.

As Minerba points out, Wolof possesses a strict phonological opposition of vowel length. This inherent characteristic gives *Wolofal* (Wolof poetry written in Arabic script) the *theoretical* capacity to organically absorb quantitative Arabic metrics:

> Very different is the state of the art of Wolof metrics. In this case, the derivation of the Wolof metrical system from the Arabic is easily proven, on both historical and linguistic grounds. Linguistically, the Wolof language distinguishes vocalic and syllabic length (Diouf, 2009: 13); therefore, it can import Arabic metrical patterns without great adaptations. It will be shown here, in fact, that Arabic metres such as *basīṭ*, *raǧaz*, *wāfir* and *mutaqārib* are employed with minimal or no changes in the corpus of Wolof poems. (Minerba, 2025, pp. 2-3)

Drawing on Minerba's linguistic premise (Minerba, 2025, pp. 122-126), because Wolof can linguistically differentiate between heavy and light syllables, it *could* naturally accommodate Arabic prosodic feet (*tafʿīla*) built upon the traditional formal units of "pegs" (*watad*) and "cords" (*sabab*).

Although I make no claim to specialized expertise in *Wolofal* poetry, a rigorous empirical methodology demands profound caution regarding such comparative extrapolations. Minerba posits that

Wolofal assimilates Arabic quantitative meters "with minimal or no changes" (Minerba, 2025, pp. 2-3). However, having observed the severe Procrustean reduction inflicted upon the Swahili syllabic grid to force it into an Arabic mold—artificially retrofitted with Western accentual mechanics by substituting the unworkable Arabic vocalic length with a Greek trochee stripped of its ancient quantity and used in its modern, stress-based sense—one is naturally compelled to cast a legitimate empirical doubt over this parallel assertion. A diffusionist hypothesis of this magnitude cannot simply be accepted on mere authority; it demands uncompromising formal modeling and rigorous statistical verification against the empirical data of the *Wolofal* corpus itself before such a direct metrical transposition onto the Wolof language can be scientifically validated.

Shifting the focus strictly back to the Swahili case, the necessity of such empirical skepticism is immediately confirmed. Indeed, as will be demonstrated in the following subsections, even though Minerba concedes that Swahili lacks the vocalic quantity required for direct Arabic metrics, he ultimately subjects the verse to yet another exogenous distortion. To force the empirical data into his comparative agenda, he uncritically absorbs Clarissa Vierke's unverified accentual *Grenzschema* (see section 2.3.1) to engineer a phantom "trochaic innovation." It is precisely this observable manipulation of the Swahili matrix that validates the profound risk of a double epistemological erasure: a methodological flattening of both West African and East African endogenous metrical systems beneath the weight of imported frameworks that the analyst simply masters better.

In the Swahili context, this structural distortion is achieved not through mere methodological deafness, but through a deliberate and systematic bypassing of the endogenous discourse. Proponents of the accentual school do not simply fail to hear the creators of Swahili verse; they actively sideline their explicit testimonies. As will be detailed in Chapters 4 and 5 of this book, the Swahili-language classical poets define their compositional rules exclusively through syllabic principles, utilizing a surgically precise lexicon to govern their architecture. Consequently, there is an absolute lexical void in the Swahili language to describe "accentual feet", "beats" or imported rhythmic inventions like the trochee or *tafʿīla*. Yet, rather than treating this glaring absence as definitive empirical evidence against the accentual hypothesis, exogenous theorists simply skip over it as an irrelevant triviality. Driven by a deeply preconceived a priori bias—the compulsion to maintain a unified comparative macro-framework

across both West and East Africa—this approach refuses to yield to the endogenous data.

This systematic flattening of two distinct African linguistic realities beneath a single exogenous Arabo-centric mold demonstrates that abstract rhetorical arguments are no longer sufficient. To dismantle this epistemological erasure, we must transition to empirical verification. While this manifesto will systematically subject the Swahili models to falsification tests on the corpus to expose this exogenous projection, this methodological imperative applies equally to Wolof: such sweeping comparative claims cannot be accepted without uncompromising empirical proof.

2.4.3. The *Ḥumaynī* Thesis: Surface Isomorphism vs. Deep Metrical Matrix

Having conceded the syllabic nature of the Swahili line, Minerba attempts to locate the Arabic influence within the stanzaic architecture and the rhyme (*vina*). He develops an elaborate and historically fascinating hypothesis tracing the complex stanzaic rhyme patterns of classical Swahili poetry directly to the vernacular Arabic *ḥumaynī* tradition of medieval Yemen.

Minerba argues that "one can trace the ancestors of the complex rhyme patterns of classical Swahili poetry in a poetic genre in modern day Yemen, known as *ḥumaynī* poetry" (Minerba, 2025, p. 205) He details the specifics of this genre, noting its divergence from classical Arabic norms:

> The *ḥumaynī* metric canon is different to classical Arabic poetry in two aspects: the strictness of its quantitative metres, which do not allow *ziḥāfas*, and the complexity of its rhyme patterns. (Minerba, 2025, p. 206)

1. Mapping the Classical Triad: The Proposed Yemeni Ancestors of Utenzi, Wimbo, and Shairi

He systematically links the Swahili metrical forms to specific Yemeni stanzaic structures. For the *utenzi* meters, Minerba defines this structural parallel:

> The *mubayyat* is a composition of quatrains (*bayt*) with a rhyme pattern *"a,a,a,a/b,b,b,a/c,c,c,a/...",* similar therefore to *tasmīṭ murabbaʿ*; its lines however can be bipartite, and consequently

> the rhyme pattern becomes *'ab,ab,ab,ab/cd,cd,cd,ab/...'* [...] The simplest form of *bayt*, without internal caesura, presents evident similarities both with *tasmīṭ murabbaʿ* within the Arabic world and with *utenzi* in the Swahili one; it is thus plausible that *utenzi* has derived from this particular form. (Minerba, 2025, pp. 206, 209)

Regarding the *shairi* meter, he posits an equally direct descent:

> In contrast, the bipartite *mubayyat* was probably the origin of the *shairi* pattern. Also here the similarities between the two structures are not negligible. Not only do the rhymes follow an alternating pattern in both cases, the first hemistichs presenting a different rhyme from the second ones; but also the two forms share the trait of repeating the rhymes of the final lines of each stanza throughout the whole poem. (Minerba, 2025, p. 210)

Finally, for the *wimbo* meter, he asserts a derivation from the *tawšīḥ* section of the Yemeni stanza:

> There are several reasons for seeing *wimbo* as a direct derivation from *tawšīḥ*: above all, the fact that both these forms are tercets; secondly, the fact that all the three lines in a stanza share the same end-rhyme and, optionally, the same internal rhymes. (Minerba, 2025, p. 211)

From an empiricist and historical perspective, acknowledging structural borrowing at the level of spatial and stanzaic layout is entirely plausible. The reality of the Indian Ocean trade networks and the well-documented migrations of Hadhrami families to Pate and Lamu (Vernet, 2005, cited in Minerba, 2025, p. 212) make cultural transmission an undeniable fact. However, one must heed the key nuances provided by historians. As Thomas Vernet establishes, while these Hadhrami migrations catalyzed a literary renewal, these migrants were rapidly Swahilized within one or two generations; the resulting poetic texts were composed entirely in the Swahili language, not in Arabic (Vernet, 2005, pp. 162–164).

2. Beyond Surface Isomorphism: Structural Counter-Evidence in Swahili Poetic Architecture

It is important to recognize that Minerba's assertions are heavily anchored in his undeniable mastery of *ḥumaynī* poetry and its highly technical prosodic jargon—a dense conceptual apparatus that can easily intimidate and overwhelm those outside the field of Arabic studies. Yet, for a specialist in Swahili poetics, this declaration of faith in a perfect Yemeni-Swahili isomorphism—pleading for nothing less than a "direct borrowing" of the meters—fails to materialize fully when confronted with the actual structural reality of the texts.

To illustrate this discrepancy, Minerba conveniently reduces the *wimbo* to a simple tercet sharing the same end-rhymes and optional internal rhymes. It is true that a superficial macroscopic resemblance exists between the Swahili *wimbo* and the Yemeni *tawšīḥ* strophic forms. However, this comparative mirage collapses entirely when we analyze the deep generative logic of their respective internal rhymes. In the *ḥumaynī* tradition, an internal rhyme functions as a regular musical punctuation; it marks the boundary of a quantitative metric foot (*taf'īla*) composed of long and short syllables, which naturally divides the line into symmetrical, rhythmically driven segments.

In stark contrast, a glance at the formal generative matrices of the *wimbo* (detailed in QRG.3) reveals an exponentially more complex and different architecture. The Swahili metrical grid is not governed by quantitative feet, but by syllabic rhymes and absolute syllabic counts (*mizani*). Consequently, a single *wimbo* line can be composed to accommodate not just one, but three or even four distinct, interlocking rhyme chains on a qualitative level. More importantly, because these Swahili rhymes (*vina*) act as formal boundaries rather than musical punctuations, they can carve the line into strictly asymmetrical blocks, as demonstrated by the hyper-fractionated matrices (see QRG. 3.4, illustrating a rigid 6+3+4+2 syllabic grid).

To classify such a hyper-sophisticated, asymmetrical, and multi-layered Swahili matrix as a "direct borrowing" from the quantitative Yemeni tercet is a highly questionable leap of logic that confuses a rhythmic marker with a formal structural boundary. While these specific Swahili matrices will be verified against the corpus later in this study, their undeniable structural presence on the manuscript page already exposes the fragility of Minerba's comparative claim. The insistence on a "direct derivation" appears to be an a priori assumption forcefully superimposed onto the data—relying on a mere

surface analogy—rather than a data-driven conclusion derived from the endogenous formal reality of the text and reproducible data-driven tests.

This profound structural autonomy is further corroborated by another endogenous *wimbo* sub-meter (*mkondo*) explicitly documented by Ibrahim Noor Shariff in his foundational treatise *Tungo Zetu* (1988). Operating through what can be formalized as the "Rupture Matrix" (for the full formal presentation and algebraic equation, see QRG. 3.2), this specific architecture introduces a radical vertical asymmetry where the syllabic parameters actively shift within the exact same stanza. The first two lines operate on a severely constricted, asymmetrical grid—such as a 2+8 or 4+8 configuration—before the final line dramatically expands into a perfectly symmetrical 8+8 structure. Because the absolute line length abruptly mutates mid-stanza, it is mathematically impossible for a continuous accentual foot—whether a creolized trochee or any other imported concept—to govern the verse. To project a direct Yemeni ancestor onto such a formally fluid and arithmetically shifting grid is a highly precarious philological extrapolation. It proves once again that Minerba's diffusionist methodology relies heavily on superficial stanzaic resemblances while ignoring the distinctly different, endogenous generative logic at work: a metric driven exclusively by formal syllabic boundaries.

To further illustrate the topological complexity of the Swahili grid and the limits of this comparative hypothesis, one need only look at the *shairi* meter, specifically the traditional 19th-century *Mashairi ya Vidato* sub-meter masterfully deployed by the Mombasa poet Muyaka bin Haji al-Ghassaniy (1776–1840). Minerba posits a direct descent of the *shairi* from the Yemeni bipartite *mubayyat*, basing his claim on a partial surface isomorphism: both forms share an alternating rhyme pattern and a repeating rhyme in the final line of the stanza. However, the *Mashairi ya Vidato* demonstrates an architectural behavior entirely alien to the *mubayyat*. In this Swahili form, the traditional static isolation of the stanza is broken to engineer a recursive structural interlocking. With the necessary exception of the initial stanza—which establishes the starting baseline—the first hemistich of the first three lines in any given stanza inherits its rhyme directly from the preceding stanza. The current stanza then acts as a progenitor for the next: it introduces a new rhyme in the second hemistich of its first three lines, which then permutes to the first hemistich position of its own fourth line. Finally, this exact same rhyme crosses the stanzaic

boundary to become the foundational first-hemistich rhyme for the first three lines of the subsequent stanza. Rather than propagating a single continuous rhyme across the entire poem, this generative movement creates a sequence of overlapping links where each rhyme chain has a strict structural range of exactly two stanzas. Consequently, any intermediate stanza is anchored to the preceding stanza by one rhyme chain, and bound to the subsequent stanza by an entirely different rhyme chain. It is precisely this overlapping, cascading structural migration that physically forms the *vidato* (the notches or rungs of a ladder) scaling the poem (for a full formal presentation of this generative matrix, see QRG. 4.3). Such a dynamic, sequential chaining proves that the Swahili grid operates not as a rigidly imported Arabic shell, but as an autonomous generative engine capable of a structural sophistication that far exceeds the simple repetitive logic of its supposed Yemeni ancestor.

Ultimately, empirical rigor demands that we distinguish between the historical inspiration for a superficial spatial blueprint and the true generative engine of the metrics. When an endogenous system proves to be far more formally fractionated and topologically complex than its supposed exogenous ancestor, insisting on a direct derivation is no longer a scientific deduction. It is the "Exo-African Postulate" operating as an unquestioned dogma.

Minerba falls into the trap of the homonymic illusion: confusing a shared architectural shell with a shared metrical engine. If a Swahili poet adopted a Yemeni *mubayyat* rhyme scheme, that structural blueprint was instantly metabolized and entirely subjugated to the invariant, pre-existing laws of Swahili phonology. The imported Arabic architecture was completely stripped of its original quantitative engine and re-engineered to run exclusively on the Swahili dual matrix of *vina* and *mizani*. Thus, acknowledging an Arabic structural inspiration for the *shape* of the stanza does not validate the importation of an Arabic *rhythm* inside the line. Once a Yemeni form is cast into the Swahili syllabic grid, it ceases to be an Arabic metric.

2.4.4. The "Trochaic Innovation": Structural Inevitability vs. Generative Intent

The deepest theoretical rupture in Minerba's framework—and the exact point where he explicitly resurrects the ghost of the accentual school—occurs when he attempts to explain the internal rhythm of the

classical Swahili meters (*utenzi*, *shairi*, and *wimbo*). In a section tellingly titled *"Utenzi, shairi, wimbo and the trochaic innovation"*, he posits that these forms actively adopted an exogenous rhythmic pulse:

> The major innovation in the history of the *utenzi*, *shairi* and *wimbo* metres is their adoption of trochaic rhythms as constitutive elements of their abstract patterns. This is evident, historically, in their progressive adoption of syllabic lengths of multiples of four. (Minerba, 2025, p. 217)

To execute this profound theoretical rupture with the voices of the Swahili-language poets and the explicit rules of their craft, Minerba aligns himself strictly with the accentual framework of Clarissa Vierke, whose model was formally presented earlier in this work (see Section 2.3.1). Rather than questioning the premise of an exogenous beat, he adopts her Procrustean methodology to forcibly project a trochaic meter onto the classical Swahili verse. This mechanical maneuver relies on weaponizing a natural phonetic property of the language—specifically its penultimate lexical stress—to artificially superimpose a rhythmic pulse that does not formally exist within the endogenous syllabic matrix. Furthermore, as will be demonstrated extensively in Chapters 4 and 5, this phantom rhythm is completely absent from the endogenous lexicon and the explicit testimony of the creators of the Swahili verse themselves. When Swahili-language master poets explain the architectural construction of their compositions, they possess no terminology to designate an accentual foot, let alone an imported classical Greek concept such as the trochee. By exploiting this mere phonetic artifact, the theory creates a loophole that permits the forceful imposition of exogenous metrical templates, effectively bypassing the absolute sovereignty of the syllabic rhymes (*vina*) and the syllable count (*mizani*).

To justify how a strictly syllabic metrics suddenly adopted a classical Greek metrical foot (the trochee), Minerba relies heavily on Roman Jakobson's concept of the *Grenzschema* (boundary schema). He quotes Jakobson to establish that a meter is an abstract ideal:

> An ideal pattern is intrinsic to the lines of a poem as a limit expression. This pattern functions as a rhythmic impulse; it is imposed on each single line as a rhythmic condition; therefore, it belongs to the essential features of the line. Whereas single

> elements of this ideal pattern are reproduced constantly throughout the lines, the expression of other elements of the same pattern constitute barely a tendency. (Jakobson, 1966, as cited in Minerba, 2025, p. 114)

Following Clarissa Vierke, Minerba embraces the idea that the 8-syllable hemistich (*kipande*) (and thus the 16-syllable line of the *shairi* meter) is perfectly bisected into a 4+4 accentual ideal, yielding a trochaic rhythm:

> Vierke (2011: 35) goes on to make an important remark: even if all the previous listed forms are possible and acceptable for a *kipande*, the ideal metric pattern in *utenzi* is that of the four-syllable foot (x x σ´ σ | x x σ´ σ), because of its symmetry and balance. She explicitly defines this form as the *'Grenzschema'* of the *utenzi*. This is the form which most directly represents the trochaic rhythm, having its two metrical stresses on odd positions and a rhythmic core at the half-length of the line. (Minerba, 2025, p. 220)

1. Drawing the Metrical Target: The "Creolized Trochee" as a Texas Sharpshooter Fallacy

Minerba then executes a breathtaking theoretical leap to explain *how* this trochaic rhythm was forged. He theorizes a "creolization" between Arabic tetrasyllabic feet (like the *rağaz*) and the archaic Swahili penultimate stress:

> The creolization of the 8 syllable line, therefore, is inspired partly by the Arabic metres based on tetrasyllabic feet, and partly by the structure of the archaic foot. The occurrence of the metrical stress on the penultimate syllable of the archaic foot, together with the Arabic inspiration of tetrasyllabicity, caused metric stresses to be on odd positions: *kutendána kutendána kutendána*. . . and therefore created a trochaic rhythm which for the first poets of the Swahili classical age could have seemed as innovative as the sequence of tetrasyllabic feet. (Minerba, 2025, p. 299)

This passage represents the ultimate ptolemaic epicycle of the accentual school of classical Swahili poetry. It constitutes a massive performative fallacy at the historical level: it confuses the inescapable phonological reality of the spoken language with a deliberate metrical

design. Minerba correctly observes that Swahili possesses a structurally fixed penultimate stress, and that when Swahili words are arranged inside an 8-syllable boundary, these natural lexical stresses frequently land on odd positions. But instead of recognizing this as an inescapable, passive phonological byproduct of the Swahili lexicon, he projects a deliberate, generative metrical *intent* onto the composer.

To establish the foundational metrical principle of classical Swahili poetry, Minerba postulates a highly convoluted process of metrical "creolization." According to his theory, this creolization occurs at the exact intersection where an imported Arabic inspiration of "tetrasyllabicity" collides with the endogenous Swahili phonological law of penultimate stress. He argues that forcing Swahili words into this Arabic tetrasyllabic boundary structurally pushes the metrical stresses into odd positions. However, Minerba's own phrasing betrays the extreme limit of this Exo-African logic. He explicitly writes that "the Arabic inspiration of tetrasyllabicity, caused metric stresses to be on odd positions: *kutendána kutendána kutendána*." A literal reading of this statement leads to a staggering linguistic absurdity: it implies that an exogenous Arabic metrical template actively dictated the stress patterns of Swahili words, effectively suggesting that an imported literary model provoked a deep phonological mutation within the Bantu language itself.

To provide a name for this resulting hybrid, Minerba—following Clarissa Vierke and the entirety of the accentual school before him—borrows its terminology from an entirely exogenous rhythmic tradition: the classical Greek trochaic foot. Yet, while the ancient Greek trochee was quantitative (a long syllable followed by a short one), its modern accentual application redefines it as strictly bisyllabic, consisting of a stressed syllable followed by an unstressed one: [- u] (see section 2.2.3). Minerba does not posit a direct evolutionary descent from Greek metrics; rather, he appropriates the term merely as a convenient explanatory label. Because the trochee inherently features a penultimate stress, he conceptualizes the Swahili outcome as a "creolized" trochee—essentially taking the bisyllabic [- u] foot and artificially appending two unstressed syllables to its front to account for his [u u - u] pattern.

Nevertheless, to disguise a passive phonological rule as an imported rhythmic design—and to go so far as to suggest that an Arabic tetrasyllabic boundary could dictate the phonological reality of a Bantu language—is a transparent theoretical sleight of hand. It forces a factitious musical intent onto the poet simply to justify an exogenous template.

To shield this exogenous projection from immediate empirical collapse, Minerba relies heavily on Clarissa Vierke's strategic framing of the *Grenzschema* (as formally presented in section 2.3.1). Vierke herself acknowledges that this 4+4 tetrasyllabic framework is merely an "ideal"—an unconscious rhythmic tendency supposedly internalized by the poets. By defining it as an unconscious impulse rather than a strict structural rule, the theory ingeniously attempts to evade falsifiability. When faced with the overwhelming reality of the Swahili corpus—where 8-syllable lines naturally break down into word groups of 3 and 5 syllables, or 2 and 6 syllables (what accentualists would count as two rhythmic "feet" or "beats")—Vierke resorts to an ad hoc methodological contrivance: she arbitrarily isolates a "syllabic core" —a trochee [- u]—that conveniently fits her desired stress pattern, and simply dismisses the pre-core syllables as metrically "irrelevant."[4]

This sleight of hand allows the model to perpetually confirm itself by ignoring all contradictory data. The fundamental point in this construction is that there must be exactly two beats per line, whether dealing with "perfect" lines that respect the 4+4 *Grenzschema* (two tetrasyllabic feet: [u u - u] · [u u - u]) or "imperfect" lines, such as the 3+5 configurations (tri- and pentasyllabic feet [u - u] · [u u u - u]). In every case, a trochaic core [- u] is guaranteed simply because of the Swahili phonological rule of the penultimate stress.

Epistemologically, this methodology operates as a textbook Texas Sharpshooter fallacy. We have already seen how the accentual school systematically confuses a passive linguistic epiphenomenon for a generative metrical cause. However, the retroactive isolation of a "syllabic core" compounds this error by introducing pure circular reasoning. By firing blindly into the text, drawing a metrical target exclusively around the naturally occurring penultimate stresses, and discarding the "missed" syllables as irrelevant, the model structurally guarantees its own validation. Because the absolute phonological law

---

[4] To clarify Vierke's formal notation: the variable x designates the metrically "irrelevant" syllables preceding the isolated trochaic nucleus (or core), denoted as $\sigma'\sigma$. In her system, σ' represents a stressed syllable and σ an unstressed one, yielding her idealized boundary schema:

($x\ x\ \sigma'\ \sigma \mid x\ x\ \sigma'\ \sigma$)

Translated into the formal notation used throughout this study—where a lexical stress is indicated by a dash [-], an unstressed syllable by a [u], and a structural concatenation by a center dot [·]—this exact same idealized tetrasyllabic pattern is formalized as:

[u u - u] · [u u - u]

of Swahili dictates that polysyllabic words inherently resolve into a [- u] pattern, locating a "trochee" near a word boundary is a structural inevitability, not a rhythmic achievement. Ultimately, this creates an unfalsifiable tautology—an academic rigged game where the theory cannot fail, simply because the underlying language has already provided the exact phonetic outcome the model demands.

2. Disentangling Endogenous Mechanics, Macroscopic Interfaces, and Performance: Introducing the 3-Layer Framework

However, even if a model is deliberately constructed to be unfalsifiable—whether shielded as an "unconscious tendency" or engineered through a Texas Sharpshooter fallacy—an empiricist cannot simply surrender the inquiry. We can still formalize these theoretical claims and force them back into the realm of rigorous scientific verification. If this purported trochaic ideal truly acts as a generative rhythmic impulse, it must leave a demonstrable footprint; it must manifest as a dominant statistical reality across the poetic corpus. Ironically, one need not look further than Minerba's own data to witness the empirical collapse of this model. On page 219, Minerba explicitly catalogs the highly variable stress combinations purportedly found in classical verse, drawing examples from Muyaka's *Kimya* (cited via Abdulaziz, 1979, p. 194).

Yet, before even analyzing his proposed stress pattern, a critical epistemological question arises: what is the acoustic source of these accents? Is this poem being naturally read aloud, is it "chanted," or is it fully sung? Minerba completely obscures this fundamental distinction.

Take the final line of his example: *Kimya kitangúsha mwánzi* that we formalize as a 6+2 framework according to Minerba's accentual scansion with exactly two feet:

[u u u u - u] · [- u]

This line contains three polysyllabic words, each carrying a penultimate lexical stress: the lexicon assigns the line three stresses (kí-mya ki-ta-ngú-sha mwá-nzi), whatever a given reading then does with them. We formalize them as follows:

[- u] · [u u - u] · [- u]

Yet Minerba's grid registers only two metrical stresses. A reading may weaken one of the three, but only a reading could show it, and Minerba does not say which reading he scans. Why, then, does one

lexical stress vanish from his grid? Because Minerba superimposes the unverified Werner/Vierke rule, the idealized two-beat *Grenzschema*, onto the text, deleting any linguistic reality that contradicts the exogenous model.

To resolve this confusion about where these accents come from and how Swahili poetry is actually oralized, we must conceptualize the classical poem through a 3-Layer Theory of Swahili Poetics, detailed in Chapter 3. This tripartite architecture allows us to classify poetic phenomena and to show where the exogenous accentual theories fail.

Layer 1: The Formal Foundation (The Syllabic Law). At the base of the system lies formal metrics proper, governed by the two axioms of the tradition: syllabic rhyme (*vina*) and the syllable count (*mizani*). This layer is formally flat: it specifies boundaries and counts, and nothing whatever about accent.

Layer 2: The Architectural Superstructure and the Comparative Interface. Emerging directly from the syllabic principles of Layer 1, Layer 2 is the macroscopic architecture of the text, where the boundaries set by rhyme (*vina*) and syllable count (*mizani*) become visible as hemistiches (sg. *kipande*, pl. *vipande*), sections (sg. *kifungu*, pl. *vifungu*), lines (sg. *mshororo*, pl. *mishororo*), stanzas (sg. *ubeti*, pl. *beti*), and complex rhyme patterns with names of their own, such as the *vidato*. Like Layer 1, it specifies nothing about accent. The endogenous lexicon of these structures is detailed in Chapter 4, and Chapter 8 (The Sovereign Grid: The Endogenous Logic of Swahili Syllabic Meter) is dedicated to their formal metrics.

This superstructure is where Emiliano Minerba legitimately situates his comparative analysis. At this macroscopic level, his erudite identification of structural resemblances with Yemeni forms is accurate and valuable, offering plausible historical connections that enrich Thomas Vernet's historical framework. The divergence arises only because, lacking a stratified analytical model, the comparative school tends to confuse consequence with cause: it mistakes a macroscopic layout, generated bottom-up by endogenous Swahili syllabic rules, for a template imported whole, top-down.

Layer 3: Melodies (*Mahadhi*) and the Performative Fallacy. The classical Swahili poem operates as a contract between the composer (*mtunzi*), who builds the grid of Layers 1 and 2, and the performer (*mghani*), who operates in Layer 3, the domain of the melodic overlay (*mahadhi*). The lexical stress of the words is fixed and deterministic; the passage to Layer 3 grants the singer wide melodic freedom. A singer may stretch vowels, suppress lexical stresses, or shift an accent

from position 2 to position 1 to syncopate the rhythm. A single line can thus receive an open range of acoustic realizations, depending on the melody the singer chooses.

It is here that Minerba, following Clarissa Vierke and Alice Werner before her, commits what this book calls the performative fallacy: they take the elastic, non-deterministic variations of a Layer 3 performance for a metrical law, and place in Layer 1 an accent that Layer 1 does not contain. To support this, Vierke sets aside the melodic freedom of the tradition by introducing a category of oralization that she terms "chanted," which she opposes to a fully "sung" performance governed by the *mahadhi*. As established in section 2.3.5, Vierke herself acknowledges that this "chanted" mode has no corresponding term in Swahili dictionaries. Yet she uses this undocumented category as the ultimate confirmation of her theory. By selecting performances in this "chanted" mode where a vocalization happens to produce her two-beat rhythm, while setting aside both the lexical structure of the words and every other performance mode that fails to confirm it, she uses a variable oralization to validate, after the fact, a metrical law posited in advance.

Accent enters the picture from a different direction, and the distinction is decisive for everything that follows. Swahili assigns one lexical stress to each phonological word, on its penultimate syllable. When a composer fills a fixed eight-syllable line or hemistich, the words chosen therefore carry a determined accentual structure, which we may compute exactly. But that structure is a property of the Swahili language, not a specification of the meter: it lies outside all three layers. A formal reading realizes it more or less faithfully; a sung performance may override it. It is here that the accentual school makes its founding move: it hears, in readings and performances, prominences produced by Bantu word structure and by the performance itself, and takes them for a metrical design.

To illustrate what the lexicon of a line delivers, consider Stanza 1 of the poem *Uwongo* by the master poet Muyaka bin Haji (Hichens, 1940, p. 92-93). Applying the penultimate rule to the words of this regular 8+8 *shairi* stanza, an operation that requires no recording and that any reader may repeat, yields the following matrices, in the notation of section 2.5.1.

L1: Karatási nakutúma (8) + léo unitumikíye (8)
Matrix: [u u - u] · [u u - u] + [- u] · [u u u u - u]

L2: Kabaleghéshe kalíma (8) + uzídi kuséma náye (8)
Matrix: [u u u - u] · [u - u] + [u - u] · [u - u] · [- u]

L3: Kudanganyána si kwéma (8) + kúlla jámbo kwa ndiáye (8)
Matrix: [u u u - u] · [u - u] + [- u] · [- u] · [u - u]

L4: Yalekéze yalekéye (8) + tus'ambiáne uwóngo (8)
Matrix: [u u - u] · [u u - u] + [u u u - u] · [u - u]

Here the lexicon itself yields unstable, non-symmetrical accentual structures. The number of accentual groups G varies across the stanza: some hemistichs yield two groups, as in L1H1, others three, as in L2H2 and L3H2. Out of eight hemistichs, the supposed 4+4 trochaic ideal, where the matrix would yield an identical [u u - u] · [u u - u] footprint, appears only twice, in L1H1 and L4H1. The rest fragments into asymmetrical combinations. Nothing here depends on how the stanza is pronounced: this is the material the composer selected, and no performance is required to produce the fragmentation.

However, concluding from a single stanza that the *Grenzschema* is invalid would be premature. This local collapse is a preview of the systematic verification conducted in Chapter 9, where the same method is applied to broader corpora. If the fragmentation observed here recurs across them, and if the 4+4 configuration occurs in poetry no more often than ordinary prose supplies it, the *Grenzschema* is falsified as a metrical principle. The remaining stress patterns, including the asymmetrical, non-4+4 configurations, are then tested for any alternative rhythmic recurrence. Chapter 9 finds none: the regularities that appear are lexical coincidences, not a generative metrical rule.

We therefore turn from argument to experimental verification, which proceeds in steps. To determine whether a bipartite tendency exists, the first step tests the lexical accentual structure of the written texts: because Layer 1 contains no accentual specification whatever, this structure must be evaluated as a phonological property of the language, independent of the meter. The second tests, by ear, what recorded formal readings realize of that structure; an instrumental measurement on a hostile sample is added for information. The lexical step shows what the words assign; the auditory step, what a reading does with it. The conditions under which each model would fail are set out in sections 2.5.1, 2.5.2 and 9.2.4; the verdict belongs to the datasets presented in Chapter 9.

2.5. Formalizing the Accentual Models for Experimental Testing

To transition from the realm of philological speculation into the domain of empirical science, we must subject the descriptive theories of the accentual school to mathematical reduction. A theoretical paradigm that relies on impressionistic descriptions, such as a "fierce beat" or an "underlying rhythmic tendency," remains inherently unfalsifiable. To test these claims against the corpus (which will be done in Chapter 9), we must first translate them into clear matrices and algebraic equations.

2.5.1. Clarissa Vierke: The Syllabic Baseline and the Accentual *Grenzschema*

Clarissa Vierke's metrical proposition for the 19th-century *utendi* is highly complex because it does not discard the endogenous syllabic rules entirely; rather, it subjugates them to an exogenous accentual core. Recognizing that the absolute limit of the eight-syllable line cannot be empirically denied, Vierke designates the classical Swahili metric as a *gemischt-quantifizierend* (mixed-quantifying) system. In this hybrid framework, the poet is supposedly bound simultaneously by the external architectural frame of the syllable count and driven by an internal, generative engine of accentual beat.

To formalize this *gemischt-quantifizierend* theory, we must establish a dual-stratified matrix: one stratum for the undisputed syllabic baseline, and a second stratum for the hypothesized accentual superstructure. Before proceeding, however, it is essential to explicitly untangle a profound theoretical inversion at the heart of this model. Vierke actively reverses the true structural hierarchy of the verse. In her framework, the endogenous syllabic count—the exact mechanism that Swahili master poets define as their true generative core—is relegated to the status of a mere peripheral shell. Conversely, she artificially elevates the accentual beat to the role of the internal driving engine. In reality, this stress is a mere epiphenomenon, a natural byproduct of Swahili phonology (where stress inherently falls on the penultimate syllable of a word) that is entirely devoid of any generative metrical function. Projecting a structural intentionality onto this basic linguistic reflex constitutes a textbook example of the Texas sharpshooter fallacy. Just as the proverbial marksman fires

randomly at a barn and only later paints a bullseye around a cluster of bullet holes to claim a deliberate hit, Vierke—providing the methodological template later uncritically adopted by Minerba—observes the natural phonological stresses that inevitably emerge when Swahili words are placed inside an eight-syllable grid. She then draws an artificial "accentual foot" around them *post facto*, falsely claiming that the poet was intentionally aiming for a rhythmic beat, rather than simply obeying the strict endogenous arithmetic of the *mizani*. To prevent any confusion between this theoretical projection and the endogenous reality, our matrices must strictly reflect this epistemological divide.

1. Stratum 1: The Syllabic Baseline

The foundational mathematical constraint, which Vierke accepts as the rigid boundary of the verse, is the absolute isometric syllable count (*mizani*) and the terminal syllabic homophony (*vina*). However, while Vierke restricts her focus to the classical eight-syllable iteration, the true endogenous model is an abstract matrix operating on a variable syllabic constant, $\boldsymbol{N}$ (where $\boldsymbol{N}$ typically equals 6, 8, or 11). As established in the exhaustive formalization of Chapter 8 (see section 8.3. The *Utenzi* Meter) and summarized in the Quick Reading Guide (see QRG.2.1), the *utenzi* stanza (*ubeti*) is defined by the following abstract parameters:

$\boldsymbol{n_{max} = 4}$: Number of lines per stanza

$\boldsymbol{H = 1}$: Number of hemistiches per line; the line is undivided)

$\boldsymbol{S = N}$: Total syllable count per line, Vierke's corpus utilizes $\mathbf{N = 8}$)

This translates into the algebraic equation of the *utenzi*:

$$\boldsymbol{Utenzi = 3(N_a//) \cdot N_X//}$$

Where $\boldsymbol{a}$ represents the variable stanzaic rhyme, $\boldsymbol{X}$ the global poem-length rhyme, and // the line boundary.

2. Stratum 2: The Accentual Superstructure

Before detailing the matrices of this accentual superstructure, it is necessary to define the specific symbolic alphabet used to formalize its variables. Since the syllabic parameters were established in Chapter 8, the following symbols are introduced exclusively for the accentual model to enable statistical counting and position tracking.

2.a. The Symbolic Alphabet

$\boldsymbol{G}$**:** The total number of accentual groups per undivided line, or per hemistich when the line is divided. Each stressed syllable defines one group: that syllable, the unstressed syllable that follows it, and any preceding unstressed syllables up to the previous group. The accentual group formalizes what the authors of the models under test call a "foot": the part of the line or hemistich that carries one "beat", their term for the stressed syllable. Since every foot carries exactly one beat, G counts both the feet and the beats of these models.

$\boldsymbol{A_i}$**:** The $\boldsymbol{i}$-th Accentual Nucleus (or rhythmic group). It is a structural set of syllables encompassing the stressed syllable, the obligatory unstressed syllable that follows it (acting as the boundary), and any preceding unstressed syllables up to the adjacent group. In the terms used in section 2.3.1 to describe Vierke's framework, it comprises the trochaic nucleus [- u] and the pre-nucleus that precedes it.

$|\boldsymbol{A_i}|$**:** The syllabic count formula. It calculates the exact number of syllables contained within the $\boldsymbol{i}$-th accentual nucleus (e.g., $|\boldsymbol{A_1}| = \boldsymbol{4}$). This is the key metric for statistically tracking foot sizes.

$\boldsymbol{P_i}$**:** The positional index of the stressed syllable that anchors the nucleus $\boldsymbol{A_i}$., ranging from 1 to the length of the relevant metrical segment (undivided line or hemistich).

$-$**:** A stressed syllable, acting as the phonetic core of the nucleus.

$\boldsymbol{u}$**:** An unstressed, weak, or neutral syllable.

[…]**:** The boundaries of the accentual nucleus ($\boldsymbol{A_i}$).

· **(Central dot):** The formal concatenation operator. It denotes the strict sequential chaining of structural units to generate a higher-level form. Its specific scope is relative to the metrical model being formalized: within the accentual model, it operates at a micro-structural level, concatenating rhythmic feet to construct the relevant metrical segment (undivided line or hemistich).

2.b. Alice Werner's Bipartite Legacy: The Axioms Defining Two Accentual Groups Per Line

Upon the rigid S=8 frame (see Stratum 1), Vierke superimposes her accentual theory (Stratum 2). She postulates that the eight-syllable line is inherently bipartite, containing exactly two stressed syllables,

or "beats", which divide the line into two "feet". The axioms below formalize this postulate in the terms of the alphabet above; the same notation serves for Minerba's model (section 2.5.2).

Axiom 1: $\boldsymbol{G = 2}$ (Every line strictly contains two metrical beats).

Axiom 2: The second beat is permanently locked onto position 7 ($\boldsymbol{P_2 = 7}$), due to the natural penultimate stress of the Swahili phonological word terminating at the line boundary.

Axiom 3: The first beat ($\boldsymbol{P_1}$) is flexible and dictated by syntax. Using $\boldsymbol{u}$ to denote an unstressed (or "weak/neutral") syllable and $-$ to denote a stressed syllable (the "rhythmic core"), we can map the matrices Vierke proposes.

2.c. The Ideal Matrix (The *Grenzschema*)

When the first stress falls on the third syllable ($\boldsymbol{P_1 = 3}$), the 8-syllable line achieves perfect mathematical symmetry. The line is divided into two symmetric 4-syllable feet. Vierke designates this as the cognitive ideal, or *Grenzschema*. To quantify both the syllabic volume of each foot and the positional alignment of the stresses for our statistical mapping, we formalize the ideal matrix as follows:

The Syllabic Count Formula (Nucleus Size):

$$|\boldsymbol{A_1}| = \boldsymbol{4}\,;\ |\boldsymbol{A_2}| = \boldsymbol{4}$$

The 8-syllable line is perfectly bisected into two 4-syllable groups.

The Positional Grid (Stress Alignment): With the primary stresses anchored at positions $\boldsymbol{P_1 = 3}$ and $\boldsymbol{P_2 = 7}$, the internal structure of the two nuclei maps out positionally as:

$$\boldsymbol{A_1 = [1\ 2\ -\ 4]}$$

$$\boldsymbol{A_2 = [5\ 6\ -\ 8]}$$

Which translates directly into Vierke's idealized accentual matrix, using the $\cdot$ operator to denote the linear concatenation of the two nuclei:

$$\boldsymbol{Matrix_{ideal} = A_1 \cdot A_2 = [u\,u\ -\ u] \cdot [u\,u\ -\ u]}$$

The Asymmetrical Matrices (Permitted Variations):

Because $\boldsymbol{P_1}$ fluctuates according to natural speech syntax, Vierke formally accepts the following asymmetrical fractionations of the 8-syllable frame. (Divisions like 7-1 or 1-7 are mathematically excluded due

to Swahili's penultimate stress rule requiring at least one unstressed syllable to follow the core).

2-6 Division ($\boldsymbol{P_1 = 1}$): $[-\,\boldsymbol{u}] \cdot [\boldsymbol{u\,u\,u\,u\,-\,u}]$

3-5 Division ($\boldsymbol{P_1 = 2}$): $[\boldsymbol{u\,-\,u}] \cdot [\boldsymbol{u\,u\,u\,-\,u}]$

5-3 Division ($\boldsymbol{P_1 = 4}$): $[\boldsymbol{u\,u\,u\,-\,u}] \cdot [\boldsymbol{u\,-\,u}]$

6-2 Division ($\boldsymbol{P_1 = 5}$): $[\boldsymbol{u\,u\,u\,u\,-\,u}] \cdot [-\,\boldsymbol{u}]$

3. Epistemological Parameters: The Conditions of Falsifiability

For Vierke's *gemischt-quantifizierend* model to hold the status of a generative metrical law, it must be testable. Under Karl Popper's epistemology, a scientific proposition must contain the exact conditions of its own falsification. If the model can infinitely absorb contradictions without breaking—if every exception is merely labeled a "stylistic liberty"—then it is not a scientific theory, but an aesthetic ideology.

By reducing her theory to the matrices above, we can establish three strict, empirical conditions for falsification to be applied during the tests of Chapter 9.

3.a. The Violation of the Two-Beat Axiom ($G \neq 2$)

The entire architecture of Vierke's theory rests on the claim that every line carries exactly two beats, hence two accentual groups (G=2), creating a bipartite tension. If a mapping of a classical *utenzi* shows that a significant share of its lines carry three or four lexical stresses, the G=2 axiom can survive only if readings suppress the surplus stresses; if formal readings realize them, the axiom is refuted. For example, if a poet builds an eight-syllable line from four disyllabic words (e.g., *Mtu huyu ana mambo*, "This person has issues"), the lexical rule assigns it four stresses, hence four accentual groups: [- u]·[- u]·[- u]·[- u]. If the words routinely give G=3 or G=4, and formal readings realize these stresses, the hypothesis of a bipartite accentual meter is falsified by the word structure of the language itself.

3.b. Testing the *Grenzschema* as a Statistical Tendency

Vierke posits the $\boldsymbol{P_1 = 3, P_2 = 7}$ symmetry as the cognitive prototype (*Grenzschema*) [u u - u]·[u u - u] that the poet allegedly holds in their mind. A cognitive prototype, by definition, must exert a

statistically significant gravitational pull on the text. If our empirical pattern analysis in Chapter 9 demonstrates that the $\boldsymbol{P_1 = 3, P_2 = 7}$ stress pattern occurs with no greater frequency than random statistical probability—meaning it appears exactly as often as one would expect by purely coincidental combinations of Swahili vocabulary—then it is not a *Grenzschema*. It cannot be categorized as a generative intent if its manifestation is indistinguishable from random lexical chance.

3.c. The Arbitrariness of Asymmetrical Distributions

The final condition of falsifiability targets the asymmetrical variations of nucleus sizes $|\boldsymbol{A_1}|$and $|\boldsymbol{A_2}|$(e.g., the 2-6, 3-5, 5-3, and 6-2 syllable distributions). In a true metrical system, variations follow identifiable periodicities or functional rules (e.g., an inverted nucleus used consistently at the beginning of a stanza for rhythmic emphasis). If the distribution of these asymmetrical nuclei ($\boldsymbol{A_i}$) across the corpus proves to be completely chaotic and entropic, entirely subordinated to the random size of the words the poet needed to fulfill the 8-syllable count (*mizani*), the accentual hypothesis is terminally wounded. If the "rhythm" is solely dictated by whichever word fits the syllabic gap, then the rhythm is merely an accidental phonetic byproduct, not a generative law. A metrical law dictates the words; if the words dictate the meter, the perceived accentual rhythm is an illusion.

2.5.2. Formalizing the Minerba Accentual Model for Experimental Testing

Any genuine return to empirical science dictates that Minerba's "trochaic innovation" in *utenzi*, *wimbo* and *shairi* meters must also undergo a strict mathematical reduction. As long as a hypothesis remains shielded by descriptive rhetoric, it is scientifically untestable. Therefore, to prepare his model for the tests of Chapter 9, we must strip away the discourse and explicitly translate his claims into both formal matrices and their corresponding algebraic equations.

Since Minerba explicitly concedes the syllabic baseline of the classical Swahili verse (2025, p. 149), we do not need to re-formalize the purely syllabic structures of the *utenzi*, *wimbo*, and *shairi* meters. For these endogenous notations, the reader is referred directly to the Quick Reference Guide and Chapter 8.

To formalize Minerba's accentual superstructure—which acts as a direct theoretical extension of Clarissa Vierke's system to encompass not only the *utenzi* meter but also the *wimbo* and *shairi* meters—we utilize the exact same symbolic alphabet defined in section 2.5.1 (e.g., G for the number of accentual groups, A_i for the nucleus, $|A_i|$ for syllabic volume, and P_i for positional indices).

1. The Minerba Matrix (The Creolized Trochaic Grenzschema*)*

Minerba's model relies on the *utenzi* octosyllabic line (S=8) and *shairi* octosyllabic hemistiches forming 16-syllables line ($S = 16$) and extends this logic to the *wimbo* line ($S = 12$, frequently divided as 4+8 or 8+4). He theorizes that these structures are multiples of a fundamental tetrasyllabic foot ($4 \times 2 = 8$, or $4 \times 3 = 12$) intended to generate an Arabic-inspired trochaic rhythm.

In Minerba's framework, this specific rhythm is the direct product of metrical creolization. According to his hypothesis, the imported tetrasyllabic templates of Yemeni Arabic poetry collided with the endogenous Swahili phonological reality—specifically, the lexical stress that falls on the penultimate syllable of a word. When a four-syllable Arabic measure is mapped onto the Swahili lexicon, this stress strikes the third position of the foot whenever those four syllables form a single phonological word, creating a strong-weak termination, that is a trochee [- u]. Minerba interprets this mechanical collision not as a mere phonological byproduct, but as a deliberate historical and cultural fusion. He argues that this cross-pollination birthed a novel, hybrid metric: the "creolized trochee" [u u - u] which supposedly reshaped the classical Swahili line into a sequence of intentional trochaic feet.

It should be noted that the collision Minerba describes does not transmit the Arabic unit but reverses it. In the standard description of Arabic prosody the rhythmic core, the *watad majmūʿ*, is rising — a short syllable followed by a long one, the accent falling on the second — whereas the nucleus he posits for Swahili is falling. The *creolized trochee* would therefore consist, in the first place, in inverting the very foot it claims to inherit, and that inversion requires no historical explanation at all: Swahili stress falls on the penultimate syllable of the phonological word, which guarantees that an unstressed syllable follows every accent, so that a rising nucleus is not available in the language. This point is taken up in section 3.3.4, where it is offered as an observation and not as a demonstration, the present author's field being Swahili studies and not Arabic prosody.

1.a. The Octosyllabic Utenzi Line ($S = 8$) and Shairi Hemistich ($H_1 = H_2 = 8$) Matrix

Much like Vierke, Minerba operates within a dual-stratified framework. Before imposing his accentual theory (Stratum 2), he does not deny the formal reality of Swahili poetry; rather, he fully accepts the endogenous syllabic baselines of these meters (Minerba, 2025, pp. 149, 151, 156) constituting the Stratum 1. To fully grasp how his model adapts them, it is essential to first recall these structural realities. They are studied in detail in Sections **8.3** (for the *utenzi*) and **8.5** (for the *shairi*), and are summarized in the condensed extracts from the Quick Reference Guides for the *utenzi* (QRG.2.1) and the *shairi* (QRG.4.2) below.

Utenzi -Standard Matrix and Algebraic Equation for One Stanza:

Formal Parameters: $n_{max} = 4, H = 1, N = S = 6$ or 8 or 11

The Matrix:

$L_1: [N, a]\ //$
$L_2: [N, a]\ //$
$L_3: [N, a]\ //$
$L_4: [N, X]\ //$

The Algebraic Equation:

$$\boldsymbol{Utenzi = 3(N_a\ //) \cdot N_X\ //}$$

Note: N represents the chosen syllabic constant for the poem: 6, 8, or 11 syllables.

Shairi - General Generative Matrix for One Stanza:

Formal Parameters: $n_{max} \geq 4, H = 2, S = N + M$ (where $S = [N] + [M]$, allowing either $N = M$ or $N \neq M$)

The Matrix:

$L_1: [N, a]\ +\ [M, b]\ //$
$L_2: [N, a]\ +\ [M, b]\ //$
$L_3: [N, a]\ +\ [M, b]\ //$
$[\ldots]$
$L_{n_{max}}: ([N, a]\ +\ [M, b])\ |$
$([N, b]\ +\ [M, X])\ |$
$([N, b]\ +\ [M, a])\ |$
$([N, X]\ +\ [M, Y])\ //$

The General Algebraic Equation for the *shairi* stanza:

$$\boldsymbol{Shairi_{general}} = (\boldsymbol{n_{max}} - \mathbf{1})(\boldsymbol{N_a} + \boldsymbol{M_b}//) \cdot \big((\boldsymbol{N_a} + \boldsymbol{M_b}) \mid (\boldsymbol{N_b} + \boldsymbol{M_X}) \mid (\boldsymbol{N_b} + \boldsymbol{M_a}) \mid (\boldsymbol{N_X} + \boldsymbol{M_Y})//\big)$$

While Minerba fully acknowledges this endogenous syllabic baseline (Stratum 1), he focuses his analysis on specific configurations that operate strictly on multiples of four syllables. To operationalize his theory, he fixes exact parameters within the general matrices. For the *utenzi*, he isolates the octosyllabic constant, setting $N = 8$ (thus $N = S = 8$). For the *shairi*, he selects the perfectly symmetrical bicolumnar configuration, setting $N = 8$ and $M = 8$ (resulting in a sixteen-syllable line where $S = N + M = 16$).

By isolating these standard octosyllabic units—whether the entire *utenzi* line or the divided hemistiches of the *shairi* verse ($H_1 = 8$ and $H_2 = 8$)—Minerba fully adopts the *Grenzschema* established by Vierke (Minerba, 2025, p. 220). Faithful to Alice Werner's bipartite legacy, this schema dictates that every octosyllabic segment must contain exactly two accentual groups. Vierke's *Grenzschema* refines this foundation by stipulating that each group should ideally be composed of precisely four syllables. Finally, Minerba adds his own layer of historical derivation, asserting that this tetrasyllabicity is a direct import from medieval Yemeni meters, resulting in a "creolized trochaic *Grenzschema*." This rhythm is structurally defined as trochaic because the primary accentual beat systematically falls on the penultimate syllable of each four-syllable sequence.

Consequently, his accentual layer matrix is identical to the Vierke model established in Section **2.5.1**, except that Minerba broadens its scope by extending this exact same eight-syllable framework to encompass the octosyllabic hemistiches forming a sixteen-syllable *shairi* line. It therefore operates on the following mathematical parameters:

The Metrical Syllabic Parameters (Stratum 1):

$S = 8$ (for the *utenzi* line), or $H_1 = H_2 = 8$ (for *shairi* hemistiches) (Establishing the fundamental eight-syllable block as the structural canvas)

The Accentual Axiom (Stratum 2):

$G = 2$ (Dictating exactly two accentual groups per eight-syllable segment)

The Syllabic Count Formula (The Tetrasyllabic Ideal):

$|A_1| = 4, |A_2| = 4$ (Defining the exact length of the metrical feet, fixing each nucleus at precisely four syllables)

The Positional Grid (Trochaic Stress Alignment):

$P_1 = 3, P_2 = 7$ (Locking the heavy metrical beats onto the penultimate syllables of each four-syllable group)

The Ideal *Grenzschema* Matrix:

$[u\,u\,-\,u] \cdot [u\,u\,-\,u]$ (Visually mapping the intended rhythm: two unstressed syllables, one strong beat, and an unstressed ending, repeated twice)

The Parity Criterion (The Extended Trochaic Rhythm): $P_1 \in \{1, 3, 5\}$

The ideal matrix above is not the whole of the claim. Minerba admits as trochaic any disposition of the two stresses that places them on odd-numbered syllables, and excludes the two dispositions that place the first stress on an even one:

> The two forms σ´ σ | x x x x σ´ σ and x x x x σ´ σ | σ´ σ can also be considered to convey a trochaic rhythm: in this case too, stresses fall on odd positions, therefore the rhythmic impression of trochee is created, even if in a weaker manner. The forms x σ´ σ | x x x σ´ σ and x x x σ´ σ | x σ´ σ are, instead, allowed variation which do not form the trochaic rhythmic impression: being closer to the archaic patterns than to the classical ones, their employment is nevertheless permitted to avoid having an over-strict reproduction of the *Grenzschema*. (Minerba, 2025, p. 220)

The model therefore operates at two degrees of strictness, and both are formalized here because both are testable. In its strict form it requires $|A_1| = 4$, $|A_2| = 4$ with $P_1 = 3$ and $P_2 = 7$, which is the *Grenzschema* proper. In its extended form it requires only that P_1 be odd, admitting the sequences 2·6, 4·4 and 6·2 as trochaic and classing 3·5 and 5·3 as permitted variations outside the rhythm. Since the second position is fixed at $P_2 = 7$ by the terminal word boundary, the extended criterion reduces to the parity of the first stressed syllable, and it is on that parity that the model is tested in Chapter 9.

1.b. The Dodecasyllabic Wimbo Matrix: The 4 + 8 Division ($S = 12$)

To fully grasp the mechanics of Minerba's dodecasyllabic adaptation, it is essential to first recall the endogenous syllabic reality of the meter. This baseline is studied in detail in Section 8.4.1. The Standard Bicolumnar Matrix and Algebraic Equation (Symmetrical and Asymmetrical), and is summarized in the condensed extract from the Quick Reference Guide (QRG.3.1) below:

Formal Parameters:

$n_{max} = 3, H = 2, S = N + M$ (allowing $N = M$ or $N \neq M$).

The Matrix:

L_1: $[N, a]$ + $[M, b]$ //
L_2: $[N, a]$ + $[M, b]$ //
L_3: $[N, a]$ + $[M, b]$ //

The Standard Algebraic Equation for the Bicolumnar *Wimbo*:

$$\boldsymbol{Wimbo_{bicolumnar} = 3(N_a + M_b\ //)}$$

Much like Vierke, Minerba operates within a dual-stratified framework. He does not deny the formal reality of Swahili poetry; rather, he fully accepts this endogenous syllabic baseline as his foundational stratum (Stratum 1), acknowledging the structural reality of the bicolumnar *wimbo* as practiced by Swahili poets and the syllabic foundation of classical Swahili poetry (Minerba, 2025, p. 149). It is only upon this empirical grid that he artificially constructs his accentual theory (Stratum 2). The present analysis will focus exclusively on how he applies this model to the asymmetrical $N = 4$ and $M = 8$ configuration.

It is upon this specific dodecasyllabic structure that Minerba extends his trochaic hypothesis. He posits that the isolated four-syllable block ($N = 4$) exists specifically to force a trochaic rhythm by isolating one tetrasyllabic sequence (Minerba, 2025, pp. 220–221). Employing the exact same symbolic alphabet introduced for the meters' modeling (utilizing the + symbol as the abstract formal marker for the metric caesura separating the hemistiches), the matrix for this configuration operates on the parameters that follow.

The Metrical Syllabic Parameters (Stratum 1):

$S = 12$, structured as a bicolumnar matrix where $N = 4$ and $M = 8$ (Establishing the endogenous twelve-syllable baseline and its specific asymmetric segmentation).

The Accentual Superstructure (Stratum 2):

$G = 1 + 2$ (Dictating exactly three accentual groups: one tetrasyllabic foot in the first colon, and two in the second colon).

The Syllabic Count Formula (Nucleus Size):

$|A_1| = 4, |A_2| = 4, |A_3| = 4$ (Defining the exact length of the metrical feet, fixing each nucleus at precisely four syllables across the entire line).

The Positional Grid: $P_1 = 3$ (relative to $N = 4$); $P_2 = 3, P_3 = 7$ (relative to $M = 8$) (Locking the heavy metrical beats onto the penultimate syllables of each respective four-syllable group, accounting for the metrical reset at the colon boundary).

The Ideal Matrix: $[u\,u\,-\,u]\,+\,[u\,u\,-\,u]\cdot[u\,u\,-\,u]$

(Visually mapping the intended rhythm, with the + symbol formally marking the internal metric caesura between the $N = 4$ and $M = 8$ hemistiches).

2. Epistemological Parameters: Falsifying the "Trochaic Innovation"

Because Minerba's framework is an expansion of Clarissa Vierke's *gemischt-quantifizierend* model, it inherits all of its structural vulnerabilities. But where Vierke confined her analysis to the *utenzi*, Minerba projects the model across three meters and superimposes on it a historical claim of imported Arabic tetrasyllabic feet. He states the extension without restriction:

> The major innovation in the history of the *utenzi*, *shairi* and *wimbo* metres is their adoption of trochaic rhythms as constitutive elements of their abstract patterns. This is evident, historically, in their progressive adoption of syllabic lengths of multiples of four. (Minerba, 2025, p. 217)

By widening the jurisdiction of the model he exposes it to points of failure that Vierke's narrower claim did not incur. Three are identified here, and each is stated in its definitive form, with its threshold and its refuting observation, among the eight conditions of falsifiability set out in section 9.2.4. What follows names them; it does not duplicate them.

The frequency of the prototype. The *Grenzschema* must occur more often in classical verse than the Swahili lexicon would produce in an eight-syllable container filled for reasons unrelated to meter, and the lines it governs must be the majority of those that carry two accentual groups. Both the strict configuration and the extended parity criterion are measured, against a prose baseline composed for the purpose. This is Condition 6.

The entropy of the accentual groups. A generative rhythm requires a unit that recurs, within the line and from one line to the next. If the number of groups G fluctuates across identical structural blocks, if the volumes $|A_i|$ fragment into sequences such as 5·3, 2·6, 3·2·3 or 2·2·4, and if the succession of configurations from line to line is indistinguishable from independent draws, no foot can be identified and no periodicity established. This is Condition 7, to which Condition 5 supplies the count.

The arithmetical limit of the tetrasyllabic foot. A segment of two or three syllables cannot host a four-syllable foot, and a hyper-fractionated configuration such as 6 + 3 + 4 + 2 admits no partition into tetrasyllabic units. Where such segments are systemically present in the corpus — and the *wimbo* matrices of Chapter 8 provide them — the derivation of the Swahili line from an imported tetrasyllabic template fails on formal grounds, before any question of prominence arises. This is Condition 8, and it is the point at which the extension to the *wimbo* becomes decisive against the model rather than merely unsupported by it.

Chapter 3. Graphic Reason, Swahili Phonology, and the 3-Layer Theory of Swahili Poetics

To accurately analyze the generative engine of the Swahili poetic act, we must permanently abandon the "Exo-African Postulate" that has long constrained a specific lineage of Western philology—namely, the accentual school. The assumption that the sophisticated metric grids of Swahili classical verse are merely localized adaptations of foreign metric systems must be replaced by a framework anchored in two fundamental, endogenous parameters: the constraints of Swahili phonology and the cognitive transformation brought about by the technology of writing.

By analyzing the intersection of these two forces, we can establish a rigorously data-driven 3-Layer Theory of Swahili poetics. This theory resolves the apparent contradictions between the formal composition of the master poets, the historical transmission of architectural blueprints across the Indian Ocean, and the highly variable acoustics of sung performance.

3.1. "*Wimbo Hutoka Ngomani*": From the Head to the Graphic Grid

Before examining the written matrices, we must first understand the cognitive genesis of the poem. Endogenous scholars have documented that the Swahili poetic intellect does not initially rely on the physical act of writing to generate its arithmetic perfection. In his foundational work *Tungo Zetu* (1988), Ibrahim Noor Shariff provides an exhaustive explanation of the cognitive origins of Swahili poetry, explicitly detailing how the ancient Swahili adage *"wimbo hutoka ngomani"* (songs come from the drum/dance) signifies that poetry organically arises from specific social events and necessities.

Shariff states:

> Swahili compositions are deeply intertwined with people's lives, and many were not composed in writing, which is why our ancestors said, 'a song comes from the dance/drum' *[wimbo*

> *hutoka ngomani]*. The word 'song' *[wimbo]* in this saying, in my thoughts, does not mean 'a composition of three lines,' but rather it means any 'composition' *[utungo]*. And the word 'dance' *[ngomani]* also does not mean 'a place where drums are beaten,' but rather it means 'where there are outcomes and a need' *[penye matokeo na haja]*. Therefore, if we interpret the meaning of this saying, we can state that its meaning is 'compositions are composed where there are outcomes and a need'. If my explanation is indeed correct, then there is no doubt that the saying tells us that Swahili compositions are usually composed merely in the head, without pen and paper. (Shariff, 1988, p. 70)[5]

Poetry is an immediate response to social necessity, a vast amount of classical Swahili poetry was historically composed entirely in the mind, without the aid of writing tools (*"hutungwa kwa kichwa tu bila ya kalamu na karatasi"*). The composer (*mtunzi*) relies on a staggering internalized arithmetic mastery, possessing the ability to calculate exact syllable counts (*mizani*) and perfectly align rhymes (*vina*) purely through cognitive processing.

This cognitive calculation is not a mechanical, pedagogical exercise. A master poet does not mechanically count syllables on their fingers like a modern school student, nor do they enumerate them consciously in their head. To understand this endogenous mastery and avoid the trap of the accentual school, one must explicitly contrast two different methods of poetic composition—a distinction extensively documented in the broader study of oral and oral-derived literatures, most notably in the Oral-Formulaic Theory pioneered by Milman Parry and Albert Lord (Lord, 1960; Parry, 1971).

In the first method—characteristic of Western accentual-syllabic poetics—the composer begins with a pre-established, abstract rhythmic matrix in their mind, for instance, a wordless trochaic rhythm: dúm-da-dúm-da-dúm-da / [- u] · [- u] · [- u].

---

[5] Swahili original : "Tungo za Kiswahili zimefungamana sana na maisha ya watu na nyingi hazikutungwa kwa maandishi, ndipo wakale wetu wakanena kuwa "wimbo hutoka ngomani." Neno "wimbo" katika msemo huu, kwa fikira zangu, halina maana ya "utungo wa mishororo mitatu," bali lina maana ya "utungo" wowote. Na neno "ngomani" pia halina maana ya "mahali panapopigwa magoma," bali lina maana ya "penye matokeo na haja". Kwa hivyo, tukifasiri maana ya msemo huu tunaweza kunena kuwa maana yake ni "tungo hutungwa penye matokeo na haja." Iwapo maelezo yangu, kweli, ni ya sawa, basi hapana shaka kuwa msemo unatueleza kuwa tungo za Kiswahili aghlabu hutungwa kwa kichwa tu bila ya kalamu na karatasi."

The poet then actively scours their lexicon for individual words whose lexical stresses satisfy this a priori metrical demand. The abstract rhythmic pulse precedes the vocabulary, and the line is dictated by a pre-existing rhythm in the head.

The Swahili classical method operates on a diametrically opposed logic. The creative engine of the Swahili poet is not "musical" in this Western sense. They do not compose by humming a pre-established abstract melody or tapping an imported rhythmic beat. Instead, the raw materials for their composition are derived directly from *ngomani*—the physical, social, and performative space where poetry, dance, and music converge. By attending and participating in *ngomani*, the poet is saturated in the rhythms, melodies, and recited verses of the living tradition (Shariff, 1988, p. 77). From this intense auditory immersion, they extract and internalize a vast repertoire of pre-constructed "building blocks": inherited formulaic half-lines, recurring syntactic structures, interlocking rhyme sets (*vina*), and stable morphological chunks that they have physically heard performed.

When composing *kwa kichwa* (in the head), the traditional master poet does not fit isolated words into a blank rhythmic tune. Rather, they operate by assembling, manipulating, and re-engineering these inherited, acoustically internalized formulaic blocks directly within their minds. Because these pre-existing building blocks already possess an inherent, stable syllabic mass, the poet relies on a deeply internalized cognitive awareness of the segmental template. Through this profound endogenous mastery, the composer develops a flawless internal balance: an instinctive formal intuition. They instantly recognize when a line is structurally "just" and perfectly fills the pre-existing cognitive void of exactly the required number of syllables (*mizani*), entirely bypassing the need to mechanically count them.

Should a segment be instinctively perceived as either deficient or excessive in its length, the classical Swahili poet never compensates by elongating or compressing a supposed exogenous "rhythm" or imported "beat." Instead, the composer executes a highly precise, surgical lexical substitution. By deliberately swapping one internalized formulaic block for another of the exact required dimensions, manipulating a morphological prefix, or inserting a specific grammatical particle, the poet ensures that the syllabic density strictly satisfies the absolute demands of the template. This structural manipulation is entirely subordinated to the semantic intent, proving the poet's absolute cognitive sovereignty over the grid.

As Kaluta Amri Abedi explicitly emphasizes on the very first page of his 1954 treatise:

> Let us remember one thing in composition, and it is this: the thoughts of the composer must generate the rhymes, not the rhymes generating the thoughts. (Abedi, 1954, p. 1)6

The poet does not mechanically slave away to fill a rhythmic slot with empty phonetic filler; rather, their profound internalized mastery allows the thought itself to organically dictate the precise morphological blocks that perfectly lock into the matrix.

Because the semantic message—the poet's underlying thoughts—remains sovereign and dictates the selection of formulaic building blocks, the flawless execution of the meter proves that this compositional process operates at a internalized cognitive level. The poet's intuitive measure is not a mechanical struggle, but rather a arithmetic and completely syllabic mastery, anchored permanently in the strict endogenous boundaries codified by Abedi. To secure the semantic intent within the formal grid, Abedi establishes the absolute structural pillars of classical poetry: *mizani* (syllable count) and *vina* (rhymes). He defines *mizani* mathematically as the precise measures (*vipimo*) used to calculate the exact length of every single line (Abedi, 1954, p. 16). Furthermore, the absolute sovereignty of these formal boundaries is enforced through the endogenous concept of *guni* (defective verse). Abedi rejects poems that fail to secure these strict metrical and rhyming locks, stating that such defective verses are simply disregarded and hold no authority within the tradition (Abedi, 1954, p. 16). Ultimately, the cognitive act of composition functions as a highly sophisticated architectural synthesis: guided primarily by their message, the poet intuitively summons and assembles deeply internalized formulaic blocks, performing a rigorous internal weighing of the verse's morphological mass against the uncompromising boundaries of a sovereign metric grid, where any failure to meet the arithmetic measure results in immediate structural rejection.

It might be objected that the composer, drawing his formulaic blocks from the acoustic reservoir of *ngomani*, thereby imports accentual shapes along with them, and that accent enters composition

---

6 Swahili original : "Neno moja tukumbuke katika utungaji, nalo ni hili ya kuwa mawazo ya mtungaji yazue vina, sio vina kuzua mawazo."

through the ear rather than through an abstract rhythmic template. The objection deserves an answer, and it has one. A formulaic block carries no accent of its own: it carries words, and the words carry the stress that Swahili assigns them, on the penultimate syllable of every phonological word. To say that the poet assembles accented blocks and to say that he assembles words are therefore two descriptions of one and the same operation.

Nor can the acoustic channel transmit a stable accentual shape. The blocks are internalized from performance, and performance is precisely where accentuation varies: the same half-line heard under two different melodies *(mahadhi)* is heard with two different accentual profiles. What survives across every performance of a block is its syllabic mass and its lexical content — which is exactly what the endogenous sources describe the composer as weighing. Shariff's account of the internalized segmental template, and Abedi's *guni*, both specify the measure and never the beat; and where a segment is felt to be deficient or excessive, the remedy documented in the tradition is a surgical lexical substitution of the exact required dimension, never the elongation or compression of a rhythmic pulse.

The objection is in any case empirical rather than doctrinal, and this is its chief interest. Any hypothesis under which composition is accentually guided — whether by an abstract template, as the accentual school proposes, or by the remembered sonority of formulaic blocks — must leave its trace in the same place: the accentual structure of the completed lines. That structure is derivable from the printed text by rule, and it is measured exhaustively in Chapter 9. The *ngomani* objection therefore does not exempt us from that measurement; it supplies a second reason to make it.

However, if the poem is born "in the head" (*kwa kichwa*), how does it achieve its monumental, multi-stanzaic architectural stability that allows it to survive across centuries? Shariff provides the key epistemological link in his section titled "*Kutunga kwa kichwa na maandishi*" ("Mental composition and the written word") (Shariff, 1988, pp. 77–80). He clarifies the relationship between the cognitive act and the physical manuscript, explaining that while the composition occurs in the mind, the subsequent transition to writing (*maandishi*) serves a highly specific, stabilizing function.

> Writing what one composes has become merely a way to preserve it for the benefit of those coming in the future days, or for those who are far away. The writing of many Swahili compositions happens after the compositions themselves have already been composed and the composer has already memorized his compositions. (Shariff, 1988, p. 78)7

It is precisely at this juncture—the transition from the cognitive calculation "in the head" to its written preservation "on the page" — that the Swahili metric system crosses the threshold into what the anthropologist Jack Goody defines as "graphic reason".

3.2. Graphic Reason and the Objectification of the Verse

The shift from the oral/cognitive space to the *ajami* manuscript is not a mere material reduplication of spoken words. It is a fundamental cognitive reorganization. Jack Goody establishes that the reduction of speech to a graphic form transforms the very nature of human cognitive operations, creating a spatial rationality that is impossible in purely oral discourse.

When the Swahili poet (*mtunzi*) commits the *vina na mizani* grid to paper, the poem is objectified. As Goody argues:

> Writing makes speech 'objective' by turning it into an object of visual as well as aural inspection; it is the shift of the receptor from ear to eye, of the producer from voice to hand. (Goody, 1977, p. 44)

This objectification is the exact epistemological arbiter of the Swahili poem. Once the verse is spatialized on the manuscript, the

7 Swahili original : "Kuandika anayoyatunga imekuwa ni njia ya kuhifadhi tu kwa manufaa ya wajao siku za mbele, au ya wale ambao wako mbali. Maandishi ya tungo nyingi za Kiswahili hutokea baada ya tungo zenyewe kwisha kutungwa na mtungaji kwisha kuhifadhi tungo zake."

fleeting, temporal nature of speech is arrested. Goody crystallizes this mechanism:

> Because when an utterance is put in writing it can be inspected in much greater detail, in its parts as well as in its whole, backwards as well as forwards, out of context as well as in its setting; in other words, it can be subjected to a quite different type of scrutiny and critique than is possible with purely verbal communication. Speech is no longer tied to an 'occasion'; it becomes timeless. Nor is it attached to a person; on paper, it becomes more abstract, more depersonalised. (Goody, 1977, p. 44)

Through this graphic reason, the Swahili poet acts as a textual architect. The page serves as a formal stabilizer that prevents the metric structure from drifting toward the natural fluctuations of speech or the temporal stretching of song. To engineer the stanza (*ubeti*), the poet relies on structural boundaries that are intrinsically literate. Goody explains the power of these graphic boundaries when discussing lists and spatial matrices:

> The list relies on discontinuity rather than continuity; it depends on physical placement, on location; it can be read in different directions, both sideways and downwards, up and down, as well as left and right; it has a clear-cut beginning and a precise end, that is, a boundary, an edge, like a piece of cloth. (Goody, 1977, p. 81)

This concept of the "boundary" or "edge" perfectly defines the Swahili *kiwango* (unrhymed caesura) and *kina* (rhyme). These are not temporal pauses for the singer to catch a breath; they are strict, visible spatial edges that lock the syllables into place (see Chapter 8). Goody confirms that "the existence of boundaries, external and internal, brings greater visibility to categories, at the same time as making them more abstract" (Goody, 1977, p. 81)

This graphic visibility allows the Swahili poet to construct the poem not as a musical score, but as a formal statement—a process Goody identifies as "graphic reductionism":

> The formula, in this strict sense, is essentially a product of graphic reductionism. In the first place the flow of speech is arrested by words or symbols being given enduring visual forms and being withdrawn from the context of the sentence. In the second, the

> linear (or rather lateral) implications of writing develop into statements of equality, or of opposition, which are also formal statements of meaning. (Goody, 1977, p. 122)

One of the particular aspects of the formula that enables us to carry out computations is the ability to retain the balance or equality between the two sides by performing the same operations on each... There is no non-visual way of doing this; the process depends upon spatial manipulation. Speech alone cannot do it; writing can. The visuo-spatial mode permits the development of a special kind of manipulation. (Goody, 1977, pp. 122-123)

It is precisely this graphic capacity to create "statements of equality" that generates the *urari wa mizani* (the balance or equality of the measure). Writing acts as a cognitive amplifier. Goody states that writing is "a tool, an amplifier, a facilitating device, of extreme importance. It encourages reflection upon and the organisation of information, quite apart from its mnemotechnic functions" (Goody, 1977, p. 109). Equipped with this spatial amplifier, Swahili theoreticians were able to develop an exhaustively precise, endogenous vocabulary completely divorced from exogenous accentual feet (see Chapter 4), proving their total mastery over the geometric arrangements of the verse.

3.3. The 3-Layer Theory of Swahili Poetics

By combining the structural implications of graphic reason with the physical realities of the Swahili language, we can resolve the tensions that have plagued Swahili metrical studies for a century. We must conceptualize the Swahili poem as a tripartite formal architecture. This 3-Layer Theory allows for a rigorous classification of poetic phenomena, proving why exogenous accentual theories are flawed.

3.3.1. Layer 1: The Formal Foundation (The Syllabic Law)

At the base of the system lies the syllabic law, the domain of formal metrics proper. It is governed by the two axioms of the tradition: rhyme (*vina*) and the syllable count (*mizani*).

As the terminological analysis of Chapter 4 details, these two pillars function as a single cognitive and formal concept, classically named

together as *vina na mizani* ("rhymes and measure"). The rhymes (*vina*) are not mere phonetic echoes at the end of a line: they are structural boundaries that cut the poetic space into segments and define them. Within these boundaries, the measure (*mizani*) fixes the number of syllables that each segment, or each line, must contain. Layer 1 thus specifies two things only: where the boundaries fall, and how many syllables lie between them. Together with the architectures of Layer 2, it belongs to the composer (*mtunzi*).

What Layer 1 requires from the language is a countable unit, and Swahili, a Bantu language, supplies it: the syllable. Swahili has no phonological opposition between long and short vowels, and a sequence of two vowels always counts as two syllables. The count is therefore determinate, and it is the only measure available: quantitative models such as the Arabic *ʿArūḍ*, which build their feet on an opposition of heavy and light syllables, have nothing to work on.

Accent does not belong to Layer 1. In Swahili, lexical stress falls on the penultimate syllable of each phonological word: it is a property of the language, not a specification of the meter, and it lies outside all three layers. The words a composer chooses give each line a lexical accentuation, which a reading realizes more or less faithfully and a performance may override (section 3.3.3, part 2). Because this stress cannot be moved within the word, composing a genuine accentual meter of alternating feet in Swahili is possible, but only at a heavy cost in vocabulary, as Chapter 10 demonstrates.

The same fixity creates the conditions of a methodological illusion. When theorists adapt Arabic metrical templates to Swahili using accentual beats, often invoking Weil's (1958) analysis of Arabic prosody, they commit a Texas Sharpshooter fallacy. Since the penultimate stress recurs in almost every word, a "rhythmic target" drawn after the fact around these stresses will always be hit; the poet is then credited with aiming at a trochee [- u], where the language alone placed the stress. The lengthening of the stressed syllable belongs to the same property of the language: it marks the stress, and it is not a metrical parameter that the poet manipulates.

Layer 1 is therefore formally flat: it specifies boundaries (*vina*) and counts (*mizani*), and nothing whatever about accent.

3.3.2. Layer 2: The Architectural Interface and the Paradox of Hyper-Sophistication

The second layer is the architectural interface: the structures that the abstract rules of Layer 1 produce once they are laid out as segments, lines and stanzas. The composer may build these structures mentally or on the page (see section 3.3.3, part 4), but only the page is accessible to analysis, and it is as a page layout that this layer is examined here. This is the domain where the history of Swahili literature, its interlocking rhyme chains, its formal stanzaic evolutions, and transoceanic borrowings manifest. Like Layer 1, it specifies nothing about accent.

It is at this layer that we must situate and evaluate Emiliano Minerba's comparative hypothesis (Minerba, 2025). Minerba proposes an elaborate and historically fascinating argument, tracing the stanzaic rhyme patterns of classical Swahili poetry (*utenzi*, *wimbo*, *shairi*) directly back to the vernacular Arabic *ḥumaynī* tradition of medieval Yemen. For example, he correctly identifies a structural isomorphism, noting that the bipartite Yemeni *mubayyat* features a rhyme pattern ("ab,ab,ab,ab/cd,cd,cd,ab/...") that closely mirrors Swahili stanzaic structure like the *shairi* meter (Minerba, 2025, pp. 206, 209).

Our data-driven models allow us to nuance the perimeter of this isomorphism. While Minerba's observation constitutes a perfect match with one specific *shairi* sub-meter—known as *Kurejelea* (Recurrence), where the final line repeats throughout the poem (see 8.5. The Shairi Meter)—it only accounts for a fraction of the meter. The vast majority of the rhyme patterns formally manifested within the *shairi* meter do not correspond to the medieval Yemeni *mubayyat*, even if we can sense a general "family resemblance" in a strictly Wittgensteinian sense.

Nevertheless, as formal empiricists, we readily concede that structural borrowing at the level of stanzaic layout is highly plausible. The Indian Ocean trade networks provide the channel for such a transmission (Vernet, 2005, pp. 162–164).

However, this historical concession raises an epistemological paradox that undermines the premise of a direct structural continuity. It dismantles the assumption that Swahili poetry is merely a passive, creolized imitation operating on the same underlying metrical engine as its supposed Arabic ancestors.

If the classical Swahili meters are simply imported Yemeni shells, we must ask a fundamental morphological question: why are the Swahili rhyming models far more complex, more finely fractionated, and topologically more sophisticated than their supposed Arabic ancestors?

Jan Knappert explicitly observed this striking divergence, noting that:

> Whereas Arabic rhyme schemes have been fixed ever since the tenth century, Swahili poets are constantly developing new patterns, and weaving ever more rhyme into their lines. (Knappert, 1979: 36)

The answer to this paradox lies in the gravitational pull of Layer 1 (the syllabic law) interacting with graphic reason. The Arabic poet, composing in the quantitative *ʿArūḍ* system, possessed a specific phonetic tool: the phonological opposition of long and short vowels. This allowed the Arabic composer to generate aesthetic complexity and rhythmic variation internally—that is, within the boundaries of a single line—by manipulating the sequence of quantitative metrical feet. [8]

The Swahili poet, working within Swahili phonology, had no such internal tool. Because Swahili has no phonological opposition of vowel length, the poet could not build quantitative feet inside the line. Faced with this phonological constraint, Swahili poets executed a brilliant structural compensation: unable to vary the metrical texture inside the line, they transferred their formal and aesthetic virtuosity to the relational architecture between the segments.

Denied the quantitative resources of the line's interior, the Swahili poet used graphic reason to hyper-sophisticate the boundaries and the linkages connecting them. The rhyme chains (*vina*) ceased to be mere phonetic ornaments and became load-bearing relational pillars, designed to lock multiple discrete segments together. As Goody notes, the visuo-spatial mode permits a special kind of manipulation. The Swahili poet transformed the stanza into a complex formal matrix,

---

[8] For an exhaustive analysis of classical Arabic metrics, see Minerba (2025, pp. 116–146). The reader is particularly encouraged to observe the magnificent combinatorial circles (dāʾiras)—designed to perfectly organize these meters—specifically located on pages 142 to 146, if only for the visual spectacle: their extreme ʿarūḍiyya technicality renders them utterly, and delightfully, incomprehensible to anyone studying Swahili poetry.

discovering that while the metrical duration of a syllable could not be altered, its relational coordinate on the graphic grid could be manipulated at will, weaving segments into highly fractionated, interlocking structures.

This is why the architectures of Layer 2, built on the rules of Layer 1, are far more complex than the Arabic models. Through the manipulation afforded by the written page, Swahili poets engineered internal inversions (*kupindua*) and cascading vertical columns (*mshuko*), which are fully formalized in the *shairi* meter section of Chapter 8, alongside hyper-fractionated asymmetrical lines (e.g., 6+3+4+2 syllabic grids) documented in the *wimbo* meter section of the same chapter. The clearest example of this Swahili architectural expansion is the *Mashairi ya Vidato*, also detailed in the *shairi* analysis of Chapter 8 (see also QRG 4.3 for a concise formal overview). In this 19th-century architecture initiated by Muyaka bin Haji al-Ghassaniy (Hichens, 1940, p. 92-93), stanzas are interlocked like the rungs of a ladder (*vidato*): the internal rhyme of one stanza migrates to become the foundational rhyme of the next. Such recursive, interlocking complexity is the work of a developed graphic reason: whether the composer resolves it mentally or on the page, it is on the page that it is laid out and can be verified.

Once we discard the "Exo-African Postulate," we are no longer bound by the false equivalence that a resemblance in surface shape implies a corresponding resemblance in deep metrical structure.

Furthermore, this epistemological break undermines the automatic reliance on crude diffusionist assumptions. A superficial likeness between some Yemeni stanzaic layouts and some Swahili meters does not constitute self-evident proof of an identical underlying mechanics. While direct historical borrowing from South Arabian models remains plausible, what Minerba (2025) identifies as a Yemeni interface was metabolized and re-engineered by the internal imperatives of Swahili phonology. The Swahili metrical grid is not a derivative; it is an autonomous, hyper-sophisticated Bantu construct.

As Lyndon Harries concluded: "Although Swahili poetry derives from Arabic poetry, in many respects there are fundamental differences. It cannot properly be compared with Arabic poetry, for it is sui generis." (Harries, 1962, p. 3)

3.3.3. Layer 3: Melodies, *Mahadhi*, and the Performative Fallacy

The final layer is the melodic overlay, the domain of the singer (*mghani*) and the performative melody (*mahadhi*). This layer provides the aesthetic flavor and the emotional resonance of the oral execution. Far from being a secondary or superficial decoration, *mahadhi* acts as a vital embellishment that beautifully elevates the absolute structural pillars of *vina na mizani*. Furthermore, this melodic environment actively participates in the cognitive process of composition. As Ibrahim Noor Shariff explicitly details, classical poets compose by internally listening to these musical frameworks, especially: "[...] those who compose by following the rhythms *[midundo]* and melodies *[mahadhi]* in their heads [...]" (Shariff, 1988, p. 77)[9]. It is from within this internal musical reservoir that the poet intuitively draws their formulaic building blocks, assembling them to precisely satisfy the absolute dimensions of the required segmental templates.

It is precisely at this deeply aesthetic layer, however, that a persistent methodological conflation occurs in exogenous metrical analyses. By approaching the poem primarily through its oral execution, the accentual school of classical Swahili poetry has systematically conflated the rhythm of the musical performance with the underlying metrical structure of the text. This specific analytical divergence is exactly what Charles O. Hartman defined as the "performative fallacy" (Hartman, 1980, p. 38).

In a striking epistemological paradox, the contemporary accentual school is theoretically aware of this trap. Emiliano Minerba, for instance, explicitly invokes Hartman's concept of the performative fallacy to warn against the "misconception of relying on rhythmic features that can potentially differ with each performance, and are not intrinsic to the text" (Minerba, 2025, p. 83). Yet, when analyzing classical Swahili meters, this vital methodological safeguard is inexplicably abandoned. By retrofitting an accentual metrical grid—such as the trochee—onto a purely syllabic matrix, Minerba and his colleagues commit the very fallacy they denounce in theory: inadvertently substituting the performative, musical embellishments inherent to Layer 3 for the invariant formal laws of Layer 1.

---

[9] Swahili original : "[...] wanaotunga kwa kufuata midundo na mahadhi kichwani mwao [...]"

1. Alice Werner's Error (1917-20) and the Accentual Legacy

To understand the devastating magnitude of this error, we must turn to Ibrahim Noor Shariff's rigorous critique of Alice Werner, one of the primary pioneers of the exogenous accentual school. In his chapter *"Upungufu unaotokana na kutokujua arudhi na mipango yake"* (Deficiencies stemming from not knowing metrics and its arrangements), Shariff dismantles Werner's 1917 essay. Werner attempted to transcribe and analyze a poem attributed to the hero Fumo Liyongo, but because she was entirely blind to the syllabic law of Layer 1 and to the architecture of Layer 2 and sought only a temporal, Western rhythm, she completely butchered the manuscript data.

Shariff clinically demonstrates that the text Werner published is a complete formal aberration:

> In her essay published in 1917, Werner transcribed for us the following composition which she accepted to be by Fumo Liyongo... If we look at the way the compositions themselves were printed, they are neither a *shairi*, nor an *utenzi*, nor a *wimbo*, nor any composition that we know. (Shariff, 1988, pp. 162-163)[10]

Because she did not understand the boundaries that the *vina na mizani* grid draws on the page, Werner took single, cohesive stanzas and arbitrarily split them into two. Shariff exposes this critical failure:

> Looking closely, we discover that Werner has divided the stanzas of the poems and printed each single stanza to make two. For example, if we mix the first two "stanzas" we get the following *shairi* stanza:... This is the first mistake relating to not knowing the metrical procedures of Swahili compositions. Therefore, relying on the writings of people who do not understand these procedures is to be increasingly led astray. (Shariff, 1988, p. 163)[11]

---

[10] Swahili original: "Katika insha yake iliyochapwa 1917, Werner ametunakilia utungo ufuatao alioukubali kuwa ni wa Fumo Liyongo... Tukitazama namna tungo zenyewe zilivyochapishwa zimekuwa si ushairi, si utenzi si wimbo wala si utungo wowote tuujuao."

[11] Swahili original : "Tukiangalia vizuri tunagundua kumbe Werner amegawa beti za mashairi na kuchapa kila ubeti mmoja na kufanya mbili. Kwa mfano, tukichanganya "beti" mbili za kwanza tunapata ubeti wa ushairi ufuatao:... Hili ni kosa la kwanza

Not only did she shatter the physical grid of the stanza, but her profound ignorance of the metrical forms led her to randomly fuse completely unrelated poems from entirely different authors. Shariff points out the absurdity:

> Third, if we investigate further, we discover, alas! Why, those stanzas are in various compositions, by various composers! There are utterances by Ali Koti, by Mwana Kupona... It shows clearly that the compositions Werner printed originate from various stanzas, by various composers, with various thoughts, which is why we saw that all the words, together, do not bring a continuous sequence of meaningful thought. There is absolutely no doubt that these compositions Werner copied down are not by Fumo Liyongo. (Shariff, 1988, pp. 164-165)[12]

He concludes his deconstruction of Werner's analytical credibility by highlighting that her failure was rooted in her inability to comprehend the metrical architectures:

> But Werner did not have knowledge of Swahili compositions and their history and their composers. It is also evident that she did not know the metrical difference between the various compositions, which is why she printed one *shairi* stanza and made it two stanzas. (Shariff, 1988, p. 165)[13]

While Werner's structural blindness at Layer 2—the surface architecture of lines and stanzas—had devastating philological consequences, effectively mutilating the manuscripts and transmitting a fractured, disjointed version of the classical artworks, her most enduring and damaging legacy occurred at Layer 3—the realm of performance and oral execution.

---

linalohusikana na kutokujua taratibu za arudhi za tungo za Kiswahili. Kwa hivyo kutegemea maandishi ya watu wasioelewa taratibu hizo ni kuzidi kupotea."

[12] Swahili original : "Tatu, tukichungua zaidi tunagundua kuwa ala! Mbona, beti hizo zimo katika tungo mbalimbali, za watungaji mbalimbali! Kuna mitongoo ya Ali Koti, ya Mwana Kupona... Inaonesha wazi kuwa tungo alizozichapa Werner zinatokana na beti mbalimbali, za watungaji mbalimbali, zenye mawazo mbalimbali, ndipo tukaona kuwa maneno yote, kwa umoja, hayaleti mfululizo wa fikira wenye maana. Hapana shaka yoyote kuwa tungo hizi alizozinakili Werner si za Fumo Liyongo."

[13] Swahili original : "Lakini Werner hakuwa na ujuzi wa tungo za Kiswahili na tarehe zake na za watungaji wake. Ni dhahiri pia kuwa hakujua tafauti ya kiarudhi baina ya tungo mbalimbali, ndipo alipochapa ubeti mmoja wa ushairi akafanya ni beti mbili."

Although she was not the absolute pioneer of the performative fallacy—Edward Steere having initiated this trajectory as early as 1870—Werner entrenched it into a dominant academic paradigm. Exhibiting a profound colonial prescriptivism, she explicitly rejected the endogenous reality of Layer 1—the foundational metrical rules and deep syllabic engineering of the text—dismissing the traditional laws defined by the Swahili poets themselves. Operating under the assumption that the exogenous observer understood the poetry better than its creators, she formally cemented the accentual fallacy in English-language scholarship by asserting:

> [...] and the conclusion I am led to adopt is that the verse is measured by beats rather than by syllables (Werner, 1917-20, p. 116)

Strikingly, this exact quote is resurrected nearly a century later by Clarissa Vierke, who uses it to open her section on "Accentuating metrics" (Vierke, 2011, p. 28). By anchoring her own methodology in Werner's foundational misjudgment, Vierke effectively perpetuates the fallacy at Layer 3.

While it is explicitly clear that she is operating within the performative realm of this third layer, her subsequent argumentation produces a profound epistemological entanglement. To expose how she violates the principle of non-contradiction and instrumentalizes filtered data to establish her theory, we must rigorously decode her treatment of her primary dataset. Analyzing the *Utendi wa Haudaji*, she states:

> I have no evidence for an *utendi* where the reciter completely adhered to a strict rhythmic pattern – which again emphasizes its status as a prototype rather than a canonic norm. From the meagre evidence that we have, we can at least presume that variation in accentuation existed in the 19th century, as Sacleux and Zache report on it. Dammann's marking of accentuation in the *"Utendi wa Haudaji"* indicates the accentuation of Kijuma. Looking at the accentual marks that he puts on monosyllabic pronouns to indicate their clitic status (see e.g. DA 159), syntax rather than strict poetic prosody seems to decide upon accentuation. (Vierke, 2011, p. 38, n. 47)

Here, Vierke's methodology collapses under a cascade of logical and epistemological errors, compounded by fundamental conceptual error in the linguistic theory she applies, thereby violating the principle of non-contradiction.

First, she explicitly acknowledges the "variation in accentuation" and the absence of a strict rhythm. This observation is entirely natural, as it accurately reflects the inherent performative variability of Layer 3. However, she immediately contradicts this premise by asserting that this accentuation is dictated by "syntax"—a fixed, deterministic property of the language, which lies outside the three layers. Logically, the accentuation cannot be simultaneously inherently variable (a performative fluctuation) and strictly invariable (locked by syntax). By attempting to merge the versatility of Layer 3 with the determinism of the language, the argument violates the basic principle of non-contradiction.

Second, she faces a theoretical imperative to resolve this logical impossibility. For the *Grenzschema*—her cognitive prototype demanding exactly two beats per 8-syllable line (see section 2.3.1) — to function as a viable metrical prototype rather than a mere fiction, it requires at least a minimum foundation of determinism and invariability; a metric law cannot rest on pure fluctuation. She attempts to establish this necessary fixity through her citation of stanza 159 (DA 159). While she explicitly introduces this stanza merely to illustrate how Dammann marks clitic pronouns, the line she refers to (*kwa mámbo bora adhímu*) conveniently features only two accents (Vierke, 2011, p. 105). This occurs because the lexical stress on the adjective *bora* is missing. Satisfied with this two-beat layout, she accounts for it by invoking "syntax", a deterministic property of the language. Yet the lexical rule assigns bora a stress, as it does every phonological word; its absence in her line records one particular realization, in which that stress was not realized, and not a property of the line. Such reductions occur even in plain reading, and syntax does not predict them (see part 2 below); here the realization was most likely a "chanted" execution. Vierke thus explains a feature of one realization by a rule that, applied to the words, predicts the opposite.

Third, Vierke commits a blatant act of cherry-picking to establish her overarching theory. She artificially freezes this contradictory, variable Layer 3 performance into a static rule simply because this specific performative anomaly—the swallowed accent on *bora*—conveniently leaves exactly two stresses (*mámbo… adhímu*). She exploits this isolated acoustic reality to illustrate the *Grenzschema*.

This theoretical construct relies on a highly restricted and manipulated sampling of performance data. To substantiate the two-beat *Grenzschema*, Vierke's protocol explicitly excludes the traditional melodic framework (*mahadhi*) of the *mghani*, focusing entirely on the "chanted" execution. Drawing from this filtered performance corpus, she isolates a statistically marginal subset of lines, retaining only those that conform to the theoretical prototype while structurally omitting lines whose lexicon yields three accentual nuclei.

This methodological vulnerability is severely compounded by a strict practice of archival gatekeeping. As explicitly acknowledged in her own citation, Vierke is fully aware that Dammann comprehensively noted the accentuation of Kijuma, and she physically possesses this complete manuscript. Yet, apart from isolated illustrations like DA 159, she actively strips the entirety of these accents from her published edition of the book.

By deliberately withholding this comprehensive phonetic dataset containing hundreds of lines, she renders her rhythmic claims entirely unfalsifiable. The reader is given no means to verify whether the unpublished accentual marks in the rest of the text actually support her two-beat *Grenzschema* or completely contradict it.

Consequently, Vierke commits a twofold error. First, she succumbs to the performative fallacy by relying on the ambiguous, variable oral execution of a single reciter (Layer 3) to dictate the formal structural laws of the poem (Layer 1). Second, by engaging in extreme cherry-picking and data filtering—isolating rare acoustic realities that fit her model while leaving aside hundreds of stanzas—she forces the Swahili line into an arbitrary accentual straitjacket, mistaking a contingent performative choice for a generative metrical intent.

2. The One-to-Many Mapping: One Line, One Lexical Accentuation, Many Performances

If a poet constructs an 8-syllable line using four disyllabic words — for example *mtu huyu ana mambo*, "this person has issues" — the phonology of Swahili assigns four lexical stresses: [- u] · [- u] · [- u] · [- u]. It is essential to specify what this structure is and what it is not. It is not a specification of Layer 1, which is silent on accent; nor is it an abstraction, since it follows by rule from the words the composer chose. It is the accentual structure that the Swahili language assigns to that

particular sequence of words, and it is computable from the printed line without any recording.

Nor is it recoverable in full from a reading. A formal reading aloud actualizes this structure only in part: it superimposes a phrase intonation under which some lexical nuclei are realized and others are not. The pattern of these reductions is not accounted for in this study, and it should not be assumed to follow the syntax: the relation between the words does not predict which nucleus is lost, and a governing word may be de-accentuated where the word depending on it is not. What is recorded here is which positions are realized, not why. This reduction occurs in ordinary reading, with no melody whatever, and it is therefore not the melodic overlay *(mahadhi)* of Layer 3. Because the reduction is real, the two must be recorded as two distinct datasets, and this study does so throughout.

In choosing our readers we do not require a match with the historical dialect of the poem: the analysis leverages the macro-Swahili penultimate stress rule, whose synchronic and diachronic stability across the dialectal spectrum and across centuries (Hyman, 2013; Nurse & Hinnebusch, 1993) makes the derivation of the lexical structure reliable for a nineteenth-century Kiamu text read by a modern speaker.

But when this exact same line transitions to Layer 3, the singer is completely free to alter this lexical accentuation to fit their musical *mahadhi*. The performer might choose to emphasize only two of those lexical stresses, or shift an accent from position 2 to position 1 to syncopate the melody.

This leads to a fundamental key concept: a single line of verse, possessing one fixed accentual structure assigned to its words by the phonology of Swahili, can generate multiple, entirely different, non-deterministic accentual placements when performed in Layer 3, depending entirely on the singer's melody. The divergence begins earlier than that, since even a plain reading aloud realizes only part of this structure; but it is in performance that it becomes unbounded, and it is performance that the accentual school has taken for the meter.

Because the accentual school operates under the performative fallacy, its methodology is inherently vulnerable to this acoustic variability. Seeking to validate their theories using Layer 3 data, early pioneers like Steere, followed by Werner, instinctively searched this highly elastic performative environment for familiar exogenous patterns, such as the trochee. Within the vast spectrum of the one-to-

many mapping, it is entirely possible—even in good faith—to isolate a specific oralization that perfectly mirrors an accentual hypothesis (for instance, a performance that happens to yield exactly two beats in an 8-syllable *utenzi* segment).

To illustrate, one could easily hypothesize a "chanted" execution of the line *mtu huyu ana mambo* where the performer suppresses the lexical stresses on positions 1 and 5, electing instead to place heavy musical beats exclusively on syllables 3 and 7. This specific acoustic realization—rendered as *mtu húyu ana mámbo*—would produce an idealized two-beat rhythm, offering a seemingly perfect materialization [u u - u]·[u u - u] of Vierke's *Grenzschema*.

However, the fatal methodological flaw lies in selecting this single acoustic iteration from among the many possible Layer 3 iterations and artificially elevating it to the status of a generative metrical law. In doing so, the accentual school systematically mistakes a contingent, transient melodic choice for the immutable structural matrix of Layer 1.

*3. The Written Contract: Bridging the Poet (*mtunzi*) and the Performer (*mghani*) in Layer 3*

The classical Swahili poem operates as a strict, uncompromising contract between the *mtunzi* (the composer and textual architect) and the *mghani* (the vocal performer). The division of labor between composer and performer can be summarized in a formula that circulates orally among readers of Swahili poetry — *shairi ni vina na mizani, mahadhi ni pambo la mghani*, "poetry is rhyme and measure, the melody is the singer's ornament." We have not been able to trace it to any printed source, and no argument in this book rests on it; it is reproduced here because it states the distinction between Layer 1 and Layer 3 more compactly than any paraphrase, and not as evidence. However, to fully grasp the mechanics of this contract, we must establish the logical imperative that binds these two roles together through the medium of graphic reason.

In the classical Swahili tradition, poetry is written specifically to be sung. The text is written down because the individual who composes the poem and the individual who sings the poem are generally not the same person. The *mtunzi* (composer) is a poet-architect, a master of both profound semantic depth and exacting structural form. He possesses the rigorous skill to construct the flawless grid of syllables (*mizani*) and interlocking rhymes (*vina*). The *mghani* (singer), on the

other hand, is a public performer, a vocal artist whose expertise lies in projecting these lines through melodies (*mahadhi*) to an audience.

Because these are two distinct individuals operating in two different artistic domains, a physical bridge connects them. The singer cannot simply guess the composer's highly complex architectural intentions. Nor can the singer rely purely on a volatile, shifting oral transmission, because a single forgotten or added syllable in spoken memory would instantly destroy the formal perfection of the grid, rendering the line defective (*guni*). Therefore, the performer relies on a stable, physical base from which to safely access the exact text. The written manuscript itself becomes this indispensable interface. The written page is the definitive medium through which the singer acquires the exact, uncorrupted words they are tasked with vocalizing.

It is precisely this transition to the written word that Lyndon Harries observed when documenting the traditional process of composition and transmission. Because the composer and the performer were distinct entities, the physical inscription of the text acted as the necessary intermediary between the architect's mind and the singer's voice. As Harries noted regarding this separation of roles, "the poet almost invariably employed a copyist or a narrator who memorized the poem and later wrote it down" (Harries, 1962, p. 5). The key point here is not whether the master poet physically held the pen themselves or dictated the lines to a scribe, but rather the graphic objectification itself. The poem materializes onto the page to lock and finalize the metrical blueprint before it can be handed over to the performer.

4. The Two Seats of the Formal Layers

One consequence of this contract deserves to be stated on its own, because it fixes the place of everything that follows. Layers 1 and 2 have two seats, and only two. They exist first in the mind of the composer, who measures the line and aligns the rhymes without pen or paper, as Shariff describes; and they exist next on the page, which fixes what he composed. The page is not a secondary copy of the mental object. It is the same structure in a form that can be inspected, transmitted and preserved, and it is the form in which the composer hands his work to the performer.

The two seats do not, however, offer the same view, and the difference falls on Layer 2. The measure of a line and the alignment of its rhymes can be held in the mind; whether a recursive architecture

spanning forty stanzas can be held there as well, no protocol can tell us. What we can observe is that it is on the page that a rhyme can be seen migrating from the interior of one stanza to the foundation of the next, as in the *mashairi ya vidato*, and that a hyper-fractionated grid can be laid out and verified column by column. This is what graphic reason adds, in Goody's sense: not the mere preservation of what was already complete, but a medium in which such structures can be built, checked and transmitted. The architectural explosion described in Section 3.3.2 is bound up with it, and the most complex Swahili forms we know are literate forms.

It may be objected that the first seat is not silent, and that the composer working *kwa kichwa* ("in the head") hears his line rather than seeing it. He does, and the objection deserves an answer, since it places sound at the very origin of the poem. But what he hears is already one way of saying the line among many: the same line imagined under another melody would sound different and would still be the same line. What is fixed in the mental seat is therefore not the sound, but what survives every way of hearing it — the boundaries *(vina)* and the measure *(mizani)*. The proof is what he transmits. He hands over a page that carries the count and the rhyme and nothing else, and the performer who receives it produces a sound the composer never heard. Had the sound belonged to what was composed, it could not have been left out of what was sent.

Layer 3 arises only at a third moment, in the voice, and it reaches the first two only through the page. Nothing that begins in the performer's mouth can travel back into them: the melody is applied to a structure that was complete before it was sung, by a poet who had no access to the performance to come.

This has a direct bearing on what can be verified and what cannot. The first seat is closed to us: no protocol reaches the formal intuition of a nineteenth-century composer. The second is open, and it is material. The experimental verification of Chapter 9 therefore operates on the page — on the measure and the rhyme boundaries that the manuscript preserves, and on nothing else that the composer determined, since nothing else is there. What a reading adds to that page comes from the language or from the reader. What a performance adds comes from the performer.

3.3.4. Locating the Scholarship

The three layers are not only a description of the poetic object; they are a diagnostic instrument. Applied to the scholarship itself, they show that the long dispute over Swahili meter has in large part been a dispute between observers describing different layers while believing themselves to be describing the same one.

The poets who have reflected on this tradition operate and build in Layer 1 and Layer 2. Kaluta Amri Abedi, Shaaban Robert, Ibrahim Noor Shariff, Ahmed Sheikh Nabhany and Mathias E. Mnyampala codify the two elements that constitute the metrical law itself: the arithmetic measure of the unit, *mizani*, and the rhyme, *vina*, understood as the boundary that carves the poetic space. That the rhyme belongs to the law and not merely to its ornament is shown by the *kifungu*, a section bounded by rhyme which carries no fixed measure at all: the boundary can operate alone, the count cannot. Upon these two elements they raise the architecture of Layer 2 — hemistiches, lines, stanzas, interlocking rhyme chains, the recursive chaining of the *mashairi ya vidato*. What their writings on the craft never specify, at either level, is a rhythm. The *guni* they define is a failure of measure or of rhyme, never of beat.

Sacleux occupies a position that precedes the question rather than answering it. He founds versification on the number of syllables and of accents alike (1909, p. 327), and he is the first to describe the accent in play with precision, calling it the oratorical accent, ordinarily falling on the penultimate of the principal words but free to displace the tonic accent, and belonging to emotionally charged delivery rather than to plain narration (1909, p. 8). What he does not possess is an operative distinction between what a composer wrote and what a reciter performs. He observes a delivery, records what he hears, and places in a single account a syllabic constant that belongs to the text and an accentual variable that belongs to its execution. This is not an error on his part; it is the state of the question in 1909, and the distinction that would separate the two had not been formulated.

Werner marks the point at which the observation hardens into a rule. She rejects the endogenous account explicitly, concluding that the verse is measured by beats rather than by syllables (1917-20, p. 116), and she fixes at two the number of those beats in the eight-syllable line. What Sacleux had described as a feature of delivery becomes, with her, a property of the meter. The step is taken without argument and without data: no mapping of the corpus accompanies

it, and the accentual notation on which it would rest is nowhere published. Vierke opens her own section on accentual metrics by quoting this very sentence (2011, p. 28), and the two-beat figure passes intact from the early 20th century (1917-20) to 2011 without any intervening verification.

Vierke, in all probability, describes Layer 3. Her own sources point that way: she works from a chanted execution, she states that she has no evidence of a reciter adhering to a strict rhythmic pattern, and she concedes that variation in accentuation existed in the nineteenth century (2011, p. 38, n. 47). She receives from Werner the figure of two beats and from Weil (1958) the mechanism that makes it intelligible, an accentual reading of the Arabic foot; what neither transmits to her is a means of establishing either. The qualification *in all probability* is required, and the reason for it is set out immediately below.

Minerba inherits this model and carries it to a different layer, which is why the placement matters. What he compares are rhyme chains as patterns: how they behave from stanza to stanza, how they permute, how they resemble the Yemeni *mubayyat*. He is not describing the rhyme as a structuring boundary but as a motif, and motifs are exactly what travels across an ocean. This study accepts the plausibility of structural borrowing at that level, and his hypothesis becomes vulnerable only where it crosses into Layer 1 — where an imported template is made to generate the measured line itself.

A remark may be added on that reading, developing the inversion signaled in section 2.5.2. It is offered as an observation and not as a demonstration, since the present author's field is Swahili studies and not Arabic prosody, in which he claims no scholarly authority. In the standard description of that discipline, the rhythmic core, the *watad majmūʿ*, is rising: a short syllable followed by a long one, the accent falling on the long. The Swahili nucleus that Vierke and Minerba posit is the reverse — falling, an accented syllable followed by an unaccented one. The trochaic innovation would therefore consist, in the first place, in inverting the very unit it claims to inherit. And that inversion requires no historical explanation at all. Swahili stress falls on the penultimate syllable of the phonological word, which guarantees that an unstressed syllable follows every accent; a rising nucleus is simply not available in the language. What is presented as a creolized innovation is the mechanical consequence of a phonological rule, described as though it were a poetic decision — which is to say that a fact belonging to no metrical layer at all has been placed in Layer 1.

The indeterminacy of the accentual source. Vierke's accentual evidence rests on the marks Dammann placed on the *Utendi wa Haudaji*, which she states indicate the accentuation of Kijuma. Nothing is recorded about the conditions under which those marks were produced. Whether Kijuma read the text, cantillated it or sang it; whether the notation reflects a single performance or several; whether Dammann transcribed what he heard or what he judged correct — none of this is documented, and none of it can now be recovered. The manuscript itself is not publicly available.

The consequence is more serious than the non-publication of the data, which is the criticism usually made and which this study has made above. Even published in full, these marks could not be assigned to a layer. An accentual notation without a documented protocol does not tell us whether it records a composed structure, a formal reading or a musical performance, and it therefore cannot ground a claim about any of them. The empirical foundation of the *Grenzschema* is indeterminate at its source, independently of any question of scholarly good faith.

This is also why the protocol of Chapter 9 documents its speaker at what may seem excessive length — her age, her region, her linguistic profile, the instruction given to her, the channel and the date of the recording. The purpose is not procedural formality. It is that an accentual annotation whose conditions of production are unknown belongs to no layer and can support no conclusion.

On the possibility of a bipartite performance mode. It may well be that certain modes of Swahili performance do organize the eight-syllable line around two salient beats, and that Vierke encountered such a mode. This study does not contest the possibility and has no means of testing it: performance melody (*mahadhi*) is Layer 3, which section 9.2.6 places outside its scope, and establishing the existence of such a mode would require an ethnomusicological corpus of multiple performances of the same text by different singers. The question is a real one and is offered to those equipped to pursue it. What can be said here is that its answer, whichever it is, would not bear on composition. A performer who reduces the accentual structure of a line to two beats is doing something to a text that was written before any performance of it existed, and by a poet who had no access to the performance to come.

3.4. Conclusion

To look for a rhythmic, accentual beat embedded in a classical Swahili poem is to look for it in the wrong place. The distinction that this chapter has established must be carried through to the end: prominence is audible in any execution of a Swahili line, and certain modes of performance may well organize it into two salient beats. What does not follow, and what the accentual school has assumed without argument, is that such a beat belongs to the meter the poet composed. The generative engine of classical Swahili metrics relies on a cognitive syllabic architecture. Working within the syllabic properties of Swahili phonology, poets compose in the mind or on the page (section 3.3.3, part 4), assembling their material into syllabic metrical grids. Once the ink meets the paper, the written text preserves this architecture and nothing else of what the poet determined. What remains is a structural blueprint laid out on the page, stripped of any melodic performance, crystallizing nothing more — and nothing less — than the arithmetical laws of syllable count and rhyming boundary. Its words carry the stresses that the language assigns them, but these belong to Swahili, not to the meter. Whatever a performer subsequently adds belongs to the performer.

The remainder of this manifesto examines whether the accentual framework can be sustained as an account of composition, along a progression of distinct and complementary axes. The ensuing chapters analyze the endogenous terminological lexicon, foreground the explicit testimonies of the poets themselves, and recover a largely forgotten scholarly legacy. On that historical and philological foundation, the demonstration advances into metrical formalization and empirical verification, under conditions of failure stated in advance for every model it examines, and closes with a poem composed to meet the accentual model on its own terms. The question at issue is not whether prominence exists in Swahili — it plainly does: the language assigns it, and a performance may reorganize it — but whether classical Swahili poetry is governed by a recurring metrical pulse, or by the law the poets themselves state: *vina na mizani* (rhyme and measure).

Chapter 4. The Terminological Proof: The Accentual Lexical Void amid the Swahili Metrical Lexicon

In cognitive or ethno-linguistics, a culture names the tools it uses to construct its art. If we examine the specialized written corpus on Swahili metrics (*arudhi*) produced by master poets and scholars, we find a technical lexicon of surgical precision, defining a formal framework governed entirely by syllabic principles.

Analyzed systematically, this endogenous terminology names two kinds of objects: structural boundaries, both rhymed (*vina*) and unrhymed (*kiwango*), and the arithmetical count of syllables (*mizani*) that these boundaries enclose. A poet may compose these structures in the mind, as Shariff reports (section 3.3.3), but the mind of the composer is closed to analysis, and this study does not claim to read it. Everything the terminology names can be checked on the page: a boundary, a rhyme, a count of syllables are all visible in the written line. Throughout this book, every claim about the meter the poets compose is therefore a claim about the page, and it is on the page that it is verified; what a reading or a performance adds to that page is examined separately (Chapter 9).

4.1. The Architecture of Compositions: An Endogenous Taxonomy

To understand the generative engine of classical Swahili poetry, we must categorize its endogenous lexicon into functional domains: the arithmetic of the measure, the structural architecture, and the boundaries marked by rhyme. This categorization reveals a sophisticated system built entirely on the Swahili syllable. Swahili poetics traditionally adopted Arabic terminology to describe its structures—loanwords fully nativized through morphological and phonological assimilation into the Swahili matrix—reflecting the historical prestige and integration of the Islamic literary tradition, without altering the Swahili nature of the underlying metrics.

To explicitly demonstrate this morphological assimilation, the standard Bantu noun class number (following Malcolm Guthrie's comparative classification) will be systematically indicated for each term in the subsequent taxonomies. This notation serves a dual empirical purpose: it provides visual proof of the loanwords' complete structural integration into the Swahili grammatical matrix, and it supplies the essential morphological data required to determine their strict syntactic agreement within the texts.

4.1.1. The Arithmetic of the Measure: The Counting Paradigm

The Swahili system completely ignores vocalic duration, syllable weight, feet and temporal beats, focusing instead on the strict numerical count of the syllables. The terminology here reflects an uncompromising arithmetic logic.

Mizani (Class 9/10) (The Measure)

Derived from the Arabic mīzān (ميزان, scale/balance) and the root *wazana* "to weigh" (Sacleux, 1939, p. 556), this term in Swahili formal metrics is not a vague rhythmic pulse. It strictly designates the absolute syllable count. To dispel any exogenous assumptions that *mizani* might refer to an accentual foot, we must refer directly to the endogenous definition provided by Ibrahim Noor Shariff: "*Mizani ni hisabu ya sauti mbalimbali katika kila neno*" ["Measure is the count of the various sounds in every word"] (Shariff, 1988, p. 42). It is the fundamental unit of measurement, completely detached from any notion of vocalic duration or stress.

Hisabu (Class 9/10) (The Count)

An Arabic loanword (ḥisāb) designating the literal act of arithmetic counting. The poet mathematically calculates the exact number of syllables to ensure the metrical contract is fulfilled according to a specific meter requirements (Shariff, 1988, p. 43).

Urari (Class 11) (The Balance)

Hypothetically borrowed from the Arabic root *'âra* (Form III *'âwara*, meaning "to bring to the same measure" or "to compare") (Sacleux, 1939, p. 968). While its common technical usage designates the balance of an account (*urari wa hisabu*) (Sacleux, 1939, p. 968),

its application in Swahili formal metrics relates directly to the overarching principle of the measure. Functioning alongside the literal arithmetical count (*hisabu*) and the absolute syllabic measure (*mizani*), *urari* indicates the specific, balanced arrangement of the syllabic measure for each individual element within the stanza, a structural harmony strictly codified as *urari wa mizani* (King'ei & Kemoli, 2001, p. 12).

Silabi (Class 9/10) (The Syllable) vs. Sauti (Class 9/10) (The Voice)

Kiswahili features a syllable structure primarily defined by an open CV sequence (a consonant followed by a vowel), a V (a single vowel), or a syllabic nasal (such as *m* or *n*). It is vital to distinguish phonology and metrical scanning from Western orthography here: a native Swahili word that appears graphically as CCV is never a true consonant cluster. It represents either a single prenasalized consonant followed by a vowel (e.g., the *mbo* in *ma-mbo*, phonologically /ᵐbo/, constituting a single CV syllable), or a syllabic nasal functioning as its own independent syllable followed by a CV block (e.g., the word *mtu*, which is strictly scanned as two distinct syllables: *m-tu*). True CCV structures, where two independent consonants share the same vowel nucleus, exist only in loanwords (e.g., the Arabic borrowing *shukrani*, Class 9/10, 'gratitude/thanks', /shu.kra.ni/; or the English borrowing *klabu*, Class 9/10, 'club/bar', /kla.bu/). There is no phonological opposition of length at the syllable level; a sequence of two vowels is always pronounced as two distinct syllables (e.g., the word *kioo*, Class 7, 'mirror', is pronounced as three distinct syllables: *ki-o-o*). This phonological reality alone immediately invalidates any application of quantitative metrics, which rely on vocalic duration.

All Swahili-language scholars utilize the concept of the syllable to analyze metrics (Abedi, 1954; Shariff, 1988; King'ei & Kemoli, 2001), establishing it as the absolute foundation of the classical grid. The syllable is the required medium and the condition of possibility for both the rhyme chains (*vina*) and the arithmetic syllabic count (*mizani*). While *silabi* is a European loanword universally adopted in modern texts, the older endogenous term *sauti* (voice) is still utilized as a strict synonym in classical scholarship to designate this quantifiable metrical unit. To prove that *sauti* functions as a discrete syllabic fraction rather than a vague acoustic phenomenon, we must

examine Ibrahim Noor Shariff's pedagogical progression. He first commands the reader to mechanically divide words to extract the metrical count:

> Let us look at my following stanza, then divide its words according to the voices contained within to get the number of measures contained in each line of the composition. (Shariff, 1988, p. 42)[14]

After breaking down a sixteen-syllable line into numbered CV blocks (*"Ha-ya ma-mbo ya ku-tu-nga... 1 2 3 4 5 6 7 8"*), he formally defines the rhyme (*kina*) by linking these "voices" to the arithmetic count:

> Rhymes are corresponding voices contained in a stanza or a section of a composition; voices which usually appear after a specific count of measures and or at the end of every line. (Shariff, 1988, p. 43)[15]

Through this clear definition, the "voice" (*sauti*) is demonstrably equated to the discrete phonetic syllable, serving as the atomic unit for both the arithmetic counting and the rhyming mechanism.

Ta[*] (Class 9/10) / U[*] (Class 11) (The Stanzaic Line Counters)

The numerical taxonomy categorizes the stanza (*ubeti*) by its absolute line (*mishororo*) count. Morphologically, this system operates on a dual structural paradigm. The Arabic loanwords uniformly apply a *Ta[*]* pattern (functioning within Class 9/10), while their endogenous Swahili equivalents consistently utilize the *U[*]* prefix (Class 11, denoting an abstract conceptual extraction).

This strict numerical progression enumerates as follows: *Tathmina* (1 line), *Tathnia* or *Uwili* (2 lines), *Tathlitha* or *Utatu* (3 lines), *Tarbia* or *Unne* (4 lines), *Takhmisa* or *Utano* (5 lines), *Tasdisa* or *Usita* (6 lines), and *Ukumi* (10 lines) (King'ei & Kemoli, 2001, p. 16).

---

[14] Swahili original : "Tuangalie ubeti wangu ufuatao, kisha tuyagawe maneno yake kwa mujibu wa sauti zilizomo ili tupate idadi ya mizani zilizomo katika kila mshororo au msitari wa utungo."

[15] Swahili original : "Vina ni sauti zilizolingana zilizomo katika ubeti au kifungu cha utungo; sauti ambazo aghlabu huonekana baada ya hisabu maalumu ya mizani na au mwisho wa kila mshororo."

These terms function strictly as arithmetic line counters rather than fixed meters definitions; a *Tarbia* (four-line) structure, for example, may manifest as either a *shairi* or an *utenzi* meter depending on internal parameters like rhyme chains and syllabic counts. Conversely, a defined metrical meter such as the *shairi* is not rigidly locked into a four-line template; it can perfectly well unfold across four, five, or six lines per stanza, as Ibrahim Noor Shariff demonstrates (Shariff, 1988; see also QRG 4.2), and as will be thoroughly analyzed in Chapter 8. This proves that these terms are strictly isolated numerical parameters—mere variables utilized to calculate and describe the broader stanzaic architecture.

4.1.2. The Structural and Cognitive Architecture

The poem is conceived as a formal structure, built geometrically on the page and sustained as a precise template in the poet's mind.

Aina (Class 9, pl. aina Class 10) (The Type / The Topological Variation)

Literally meaning "sort" or "type," *aina* designates the specific topological variations occurring within an overarching poetic meter (*bahari*). From a strictly formal perspective, a specific type of poetry (*aina*) is defined within its meter by the precise motif generated by the spatial trajectories and paths of the rhyme chains (*vina*) that structure its stanzas (*beti*).

Bahari (Class 9, pl. bahari Class 10) (The Meter / The Sea)

This nominal is a loanword derived from the Arabo-Persian root for "sea" or "ocean" and "poetic meter" (*baḥr*). Within Swahili formal poetics, it operates under two distinct but interconnected technical meanings. Neither of these meanings should be confused with the loose, thematic Western concept of a literary "genre." Instead, *bahari* strictly designates the meter and its absolute structural components.

Definition 1: The Overarching Meter

A *bahari* is the primary metrical framework of classical Swahili expression. This meter is deduced directly from the endogenous formal principles articulated by the Swahili-language master poets themselves. At the scale of the stanza (*ubeti*), the meter corresponds

to exact, determined values that constitute the cognitive and structural grid of the poem. This endogenous grid strictly defines the number of structural lines (*mishororo*) per stanza, their internal division into segments (*vipande*), and the absolute arithmetic syllable count (*mizani*) required for each block. In this framework, the poetic space is actively carved out by the specific number, phonetic nature, and spatial orientation of the rhyme chains (*vina*) on the written page, alongside any unrhymed metric boundaries (*viwango*) utilized by the poet. Ultimately, these rhymes and caesuras form the architectural blueprint of the verse, which the strict arithmetic constraint of the syllables (*mizani*) simultaneously locks into a perfect formal equilibrium.

Definition 2: The Omni-Stanzaic Rhyme and the "Habari" Shift

In a more specialized technical sense, *bahari* also designates a highly specific type of terminal rhyme. The *bahari* rhyme is an inter-stanzaic and omni-stanzaic anchor that strictly unites all the final syllables of the final segments of the final lines across the entirety of a poem. The continuous, unbroken trajectory of this rhyme chain is the absolute defining characteristic of the *utenzi* meter (though it may structurally occur in other meters).

This dual technical meaning leads to a fascinating historical and terminological nuance involving Shaaban Robert. As noted in his foundational taxonomy, Shaaban Robert conceptualized the continuous terminal rhyme of the *utenzi* not as the *bahari*, but as the *habari* (the news, the narrative, or the core message) (Robert, 1972). However, in strict formal taxonomy, as explicitly codified by theoreticians like Kaluta Amri Abedi (1954), this specific terminal rhyme is categorized as the *bahari*.

It is highly probable that Shaaban Robert's use of *habari* represents a poetic reinterpretation—or an elegant terminological inversion—of the strict metric term *bahari*. While Shaaban Robert's supreme mastery of Swahili letters is undisputed, strict formalists such as Ibrahim Noor Shariff (1988) have systematically advocated for an precise *arudhi* (metrics), occasionally highlighting the theoretical or terminological imprecisions of earlier generations of scholars. In this light, the transformation of *bahari* (the formal metric ocean/anchor) into *habari* (the narrative message) perfectly illustrates the occasional tension between rigorous structural taxonomy and the semantic intuition of a master poet.

Kifungu (Class 7, pl. vifungu Class 8) (The Section)

Literally translating to a "chapter," "part," or "section," this nominal of Bantu origin derives from the verbal root *-funga* (meaning "to close," "to tie," or "to bind"). Within the endogenous terminology, it designates a specific sequence of syllables bounded between two metric stops—whether those stops are operated by rhyme chains (*vina*) or unrhymed boundaries (*viwango*). The defining characteristic of a *kifungu*, however, is its arithmetic irregularity. Simply put, it is a section carved out by rhymes or unrhymed boundaries, but which inherently lacks a defined or constant syllabic measure (*mizani*).

To understand its exact nature, it must be strictly contrasted with the measured elements of the poem. When the syllables bounded between the rhymes or caesuras display a strict, calculated arithmetic regularity (e.g., exactly eight syllables), the sequence constitutes a hemistich (*kipande*) or a full line (*mshororo*). It is only when this arithmetic regularity fails or is absent that the sequence remains categorized merely as a *kifungu*.

Ibrahim Noor Shariff explicitly utilizes this term when mapping the architectural scope of the rhyme, placing the *kifungu* alongside the stanza (*ubeti*) as a valid structural container closed by matching sounds: "Rhymes are corresponding sounds contained in a stanza or a section of a composition [...]" (Shariff, 1988: 43)[16].

By using the term *kifungu*, the Swahili theoretical framework acknowledges that a text can be bounded by rhymes without necessarily conforming to a strict syllabic count, separating the concept of the rhyming section from the arithmetically measured line.

Kipande (Class 7, pl. vipande Class 8) (The Segment / The Hemistich)

This nominal of Bantu origin, literally meaning "piece" or "part", designates a formally balanced, arithmetic fraction of the structural line (*mshororo*). In the Swahili formal grid, the lines of a stanza (*ubeti*) can be divided into segments of strictly regular syllabic sizes under the cleaving effect of the rhyme chains (*vina*) or predefined unrhymed limits (*viwango*).

[16] Swahili original : "Vina ni sauti zilizolingana zilizomo katika ubeti au kifungu cha utungo [...]"

It is important to clarify the Western translation of this term. The Greek-derived word "hemistich" (from *hēmi-*, "half", and *stikhos*, "line" or "verse") strictly implies a bisection into exactly two parts. When a Swahili line is divided into two segments of equal syllabic weight, these *vipande* indeed function perfectly as true hemistiches. However, because *kipande* inherently means "part," the Swahili formal system is structurally broader. A single line can consist of one undivided *kipande*, two *vipande*, or be fractionated into three or four distinct *vipande*.

Regardless of the number of divisions, the use of this term implies a deliberate, formal division governed entirely by the absolute arithmetic laws of the syllabic measure (*mizani*). It is a structural division rooted in cognitive precision, ensuring that each part perfectly fits the formal grid designed by the composer.

Mkondo (Class 3, pl. mikondo Class 4) (The Sub-Meter / The Syllabic Current)

Mkondo is a purely Bantu nominal that literally translates to "current" or "flow" (such as a watercourse or an ocean current). Within Swahili formal poetics, it is utilized as a highly specific technical term to describe the strict internal syllabic subdivisions and formal flows operating within an overarching meter (*bahari*).

Structurally, a *mkondo* is defined by its precise syllabic architecture within a given meter. It corresponds directly to the exact number of segments (*vipande*) per structural line (*mshororo*) and the absolute syllabic size of these segments (the syllable count, or *mizani*). Therefore, the *mkondo* operates at the vital intersection of two foundational formal principles: it relies simultaneously on the spatial delimitation of the verse's internal parts (dictated by the generative force of the rhyme chains) and the rigorous arithmetic measure applied to those defined segments.

The terminological relationship between the meter (*bahari*) and its sub-meters (*mikondo*) reveals a profound, elegant endogenous metaphor embedded within Swahili literary theory. Just as the vast, overarching ocean (*bahari*) is physically navigated by recognizing and following its underlying, dynamic currents (*mikondo*), a primary metrical framework is structurally animated by the specific syllabic flows running through its stanzas. A single poetic sea can accommodate various distinct, internal syllabic currents without losing its absolute macro-structural identity. Thus, to compose within

a specific classical meter is to navigate its waters, and to select a *mkondo* is to deliberately channel a precise, arithmetic current within that broader ocean.

Mshororo (Class 3, pl. mishororo Class 4) vs. Mstari (Class 3, pl. mistari Class 4) — The Structural Line vs. The Graphic Line

To accurately map the Swahili metrical grid, one must rigorously distinguish between the structural line of the poem and its physical manifestation on the page. To avoid the ambiguity of the English word "verse" (which can refer to either a stanza or a line), this framework differentiates between the "structural line" and the "graphic line."

Mshororo designates the true structural line—a cognitive and structural unit within a stanza (*ubeti*) that is formally delimited by rhyme chains (*vina*) or predefined unrhymed limits (*viwango*). Depending on the specific meter's architecture, these rhyme chains may dictate both internal divisions (*vipande*, hemistiches) and the terminal boundary of the structural line. The etymological status of this term is uniquely complex: standard dictionaries note no direct Bantu root. While it shares a phonological affinity with the Arabic *miṣrā'* (pl. *maṣārī'*) and the classical Persian *mesrāʕ*—terms that specifically designate a hemistich in Arabo-Persian prosody—lexicographical sources do not establish a formal etymological link. Postulating a direct loan is highly problematic, as it would imply a profound formal category shift: taking a foreign term for a "half-line" and repurposing it to designate a complete, autonomous Swahili structural line.

Mstari, conversely, is an explicit Arabic loanword designating a "line" in a general, physical sense—such as a line drawn with a ruler, a military front line, or a line of printed text. Consequently, *mstari* is intrinsically associated with the visual and graphic order on the page, whereas *mshororo* remains tethered exclusively to the endogenous poetic structure.

As Ibrahim Noor Shariff's framework clarifies, understanding this distinction is critical: the *mshororo* is the invariant structural line defined by the formal law, but its graphic layout as a *mstari* can vary considerably. A single *mshororo* (the complete structural line formally locked by its rhyme chains) can be visually shattered across multiple *mistari* (vertical graphic lines) in modern romanized printing, whereas it traditionally unfolds as a single, continuous horizontal *mstari* in the graphic reason of Ajami manuscript culture.

Ubeti (Class 11, pl. beti Class 10) (The Stanza)

This is the complex metric unit of the poem. It is a loanword derived from the Arabic *bayt* (بَيْتٌ), which literally translates to a "house," a "tent," or a single verse in Arabic quantitative poetry (Johnson, 1939, p. 485; Sacleux, 1939, p. 104). In Swahili, this term underwent a profound architectural redefinition: *ubeti* no longer designates a single line, but a fixed, multi-line stanza. This semantic shift is beautifully illuminated by the term's structural integration into the Bantu noun class system, specifically the Class 11/10 pairing. In Swahili morphology, this specific class pair frequently operates on a singulative/collective dynamic (as documented in standard reference grammars like Ashton, 1962, and Mpiranya, 2015). Just as the language utilizes the Class 11 prefix to extract a single hair (*unywele*) from a homogeneous plural mass (*nywele*, Class 10), it conceptually treats the continuous sequence of the poem as a plural mass of structural units (*beti*, Class 10). Applying the Class 11 prefix (*U-*) thus operates a precise morphological extraction, isolating one singular, structurally cohesive architectural block (*ubeti*) from the broader sequence.

The *ubeti* is structured by syllabic rhyme chains (*vina*) that present distinctive parameters—such as their number, orientation, trajectory, and scope—which define the various poetic meters (*bahari*). These parameters delineate the lines (*mishororo*) and, where applicable, the internal segments (*vipande*) within the stanza. Consequently, the *ubeti* serves as a complex metric unit composed of a specific number of lines, predetermined according to the overarching meter. The number of segments per line and their syllabic counts can vary, defining specific metric currents (*mikondo*). A line within the stanza may contain more than two sections; however, when it consists of exactly two segments of equal syllabic weight, the *ubeti* presents a succession of lines perfectly bisected into equal hemistiches.

Tungo (Class 9, pl. tungo Class 10) (The Composition) and Diwani (Class 9, pl. diwani Class 10) (The Anthology)

Tungo is a nominal of Bantu origin derived from the verbal root -*tunga* (to compose/to string together like pearls), emphasizing the act of carefully assembling independent, equal parts into a cohesive formal whole. In its broadest sense, it designates any written composition in general, whether it be an essay, a dissertation, or a

poem. Within the specific context of formal poetics, however, Swahili theoreticians employ the term *tungo* to strictly disambiguate the discussion of poetic works, avoiding the inherent semantic overlap of the term *shairi*, which designates both a highly specific formal meter and "poetry" in a generalized sense.

Conversely, *Diwani* is an Arabic loanword (*dīwān*) repurposed to designate a complete, curated anthology of a master poet's work, functioning as a book of poetry. The etymological roots of the word brilliantly reflect the compilation process itself. According to the *Encyclopédie de l'Islam* (1961, Tome II, pp. 332-333), there are two competing and contradictory historical origins for the Arabic *dīwān* in the primary sources: either the Persian *dēv* ("mad" or "devil," a term originally applied to administrative secretaries), or the Arabic verbal root *dawwana* ("to collect" or "to register," yielding the meaning "a collection of pieces or leaves").

It is this second meaning that perfectly aligns with Swahili literary practice. A master poet does not typically sit down with the initial intent to write a *diwani* as a single, continuous project; rather, in strict accordance with the *dawwana* root ("collection of pieces"), the author actively gathers and selects individual poems they have already composed over the long term, spanning various times, locations, and literary projects.

In the world of Swahili-expression poets, this anthology is traditionally intended to offer a masterpiece or a defining legacy to the readership. The foundational model for this tradition was inaugurated in the early 19th century on the coast of present-day Kenya with the *Diwani* of Muyaka bin Haji al-Ghassaniy (1776–1840) (Hichens, 1940). Muyaka's compilation is widely known by his successors and is considered an absolute classical benchmark and a fundamental model.

4.2. The Metric Locks, Boundaries, and Trajectories

The Swahili metrical grid is secured by formal syllabic walls and by the rhymes that lock them. The terminology describes these boundaries and the geometric trajectories that connect them, all visible on the page: it presents the grid as an architecture.

Kina (Class 7, pl. vina Class 8) (The Rhyme)

The rhyme is a structural phenomenon of syllabic homophony. The etymology of this term is rich and has been rigorously documented by Charles Sacleux (1939), revealing a dual linguistic heritage. On the one hand, it is categorized as a term of ancient Bantu origin, stemming from the root *mwina* (meaning a hole, a pit, or the depths of the water) (Sacleux, 1939, p. 650). On the other hand, it is also identified as a nominal of Arabic origin, derived from *kân wa kân*, meaning the "meter of a verse (that is scanned)" (Sacleux, 1939, p. 556).

Technically, the rhyme is defined by a relation of strict syllabic homophony between the syllables of segments (*vipande*) or lines (*mishororo*). Far from being limited to the end of the line, this homophony operates at various depths within the formal grid, defining internal rhymes (*vina vya kati*), median rhymes (*vina vya katikati*), and final rhymes (*vina vya mwisho*). This homophony functions at the scale of a single stanza (*ubeti*), multiple stanzas, or the entirety of the poem.

The rhyme functions as the structural and auditory lock of the *mizani* measure. To understand exactly how the rhyme is formally and structurally interdependent with the syllable count, we must examine Ibrahim Noor Shariff's extensive definition:

> Rhymes are corresponding sounds contained in a stanza or a section of a composition; sounds which usually appear after a specific count of measures [syllables] and or at the end of every line. (Shariff, 1988, p. 43)[17]

Shariff's definition is epistemologically devastating for the accentual school: it proves that the "sound" (*sauti*) only qualifies as a rhyme when it appears after a specific, calculated arithmetic count of syllables (*baada ya hisabu maalumu ya mizani*). The rhyme confirms—both on the page and in the "mental ear" of the poet—that the arithmetic syllable count has been executed perfectly.

[17] Swahili original : "Vina ni sauti zilizolingana zilizomo katika ubeti au kifungu cha utungo; sauti ambazo aghlabu huonekana baada ya hisabu maalumu ya mizani na au mwisho wa kila mshororo."

Kiwango (Class 7, pl. viwango Class 8) (The Quantity / The Limit)

Literally translating to "quantity," "amount," or "rate," *kiwango* is a nominal of Bantu origin. Within the formal grid of classical Swahili metrics, it designates an unrhymed metric boundary or caesura.

When a structural rhyme (*kina*) is absent, the *kiwango* operates as the exact substitution for the rhyme's delimiting function. It cleaves segments (*vipande*) or lines (*mishororo*), and just like rhymes, these unrhymed boundaries can associate to form chains across the stanzas.

To understand its rigorous arithmetic nature, we must examine Ibrahim Noor Shariff's definitive explanation:

> A *kiwango* [caesura/limit] is a stop after a specific measure [syllable count] in every line of a stanza which does not have matching sounds or rhyme. (Shariff, 1988, p. 43)[18]

This definition is essential because it permanently invalidates the exogenous assumption that a caesura acts as a fluid musical rest or a zone of rhythmic elasticity utilized to compensate for missing syllables. As Shariff demonstrates through his analysis of the 11-syllable line in the *Inkishafi* epic, the *kiwango* is an immovable, arithmetic wall fixed at an exact syllabic coordinate (for instance, locking exactly after the sixth syllable: *mizani ya sita*) (Shariff, 1988, p. 55)

Building upon this strict framework, we formalized the theoretical hypothesis that the *kiwango* operates as a "rhyme zero" [*rime zéro*] (Roy, 2013b, pp. 305, 313, 1110). Because it depends entirely on the foundational "rhyme principle" to formally delimit the verse, it acts as a spatial and arithmetic lock where the terminating syllable simply possesses an undefined phonological value instead of a homophonic one. The concept of the "rhyme zero" demonstrates that the structural capacity to cleave the poetic space is absolute, relying completely on the invisible arithmetic wall of the *kiwango* even in the absence of sound repetition.

[18] Swahili original : "Kiwango ni kituo baada ya mizani maalumu katika kila mshororo wa ubeti ambacho hakina sauti za milingano au kina."

Kituo (Class 7, pl. vituo Class 8) vs. Kibwagizo (Class 7, pl. vibwagizo Class 8) (The Stop vs. The Refrain)

These nominals of Bantu origin function as crucial structural markers representing the resolution of the verse. *Kituo* derives from the verbal root *-tua* (meaning "to descend," "to settle," or "to put down"), while *kibwagizo* is rooted in the verb *-bwaga* (meaning "to throw down," "to let fall," or "to drop a burden"). Within the precise terminology of Swahili formal poetics, *kituo* operates under two distinct, albeit related, technical meanings, the second of which intersects directly with the functional concept of *kibwagizo*.

Definition 1: Kituo as the Metrical Pause or Caesura

Endogenous Swahili theoreticians utilize *kituo* to designate a metrical stop or caesura. It is a structural operation resulting directly from the generative effect of the rhyme chains (*vina*) or the unrhymed metric boundaries (*viwango*) and the syllable count (*mizani*). Depending on its spatial depth within the poet's cognitive grid, this stop actively separates internal segments (*vipande*), structural lines (*mishororo*), or entire stanzas (*beti*).

Definition 2: Kituo vs. Kibwagizo (The Terminal Line vs. The Refrain)

In a second, highly specialized technical sense, *kituo* designates the final structural line (*mshororo wa mwisho*) of a stanza (*ubeti*). This structural usage is explicitly documented by foundational theoreticians such as Kaluta Amri Abedi (Abedi, 1954, p. 19). It is further rigorously codified in modern lexicography by K. W. Wamitila in the *Kamusi ya Ushairi*, which establishes the structural synonymy between the two terms, defining *kituo* exactly as: *"mshororo wa mwisho katika ubeti wa shairi. Pia kibwagizo"* ["the final line in the stanza of a poem. Also a refrain"] (Wamitila, 2006, p. 75).

While *kituo* designates the architectural reality of this final line—functioning as the heavy structural "stop" or landing point before the stanza's syllabic measure (*mizani*) resets—*kibwagizo* emphasizes its functional and textual repetition as a refrain across the stanzas. Because it is the ultimate sequence of the stanza, this terminal line frequently carries the continuous, omni-stanzaic rhyme (such as the *bahari* in the *utenzi* meter) and functions as the definitive formal and narrative anchor of the poem.

4.2.1. The Topological Motifs of Rhyme Chains

The endogenous terminology of Swahili poetics extends far beyond the foundational pillars of the syllable count (*mizani*) and the rhyme (*kina*, pl. *vina*). The master poets of the Swahili language, who wrote about the rules of their own art, developed a highly specialized vocabulary to describe the exact topological trajectories, permutations, and structural motifs formed on the page by the interlocking rhyme chains across the stanza.

To fully grasp the profound reality of this endogenous meta poetical framework, this section incorporates the contribution of Charles M. Mnyampala, the second son of the master poet Mathias E. Mnyampala. While our terminological analysis systematically cross-references multiple foundational texts to ensure lexicographical rigor, his introductory synthesis serves as an exceptionally pedagogical entry point. It provides unassailable proof of how the Swahili poetic tradition endogenously names, describes, and categorizes its formal rhyme patterns. By explicitly detailing these structural motifs and providing direct, concrete examples within an open-access text (Mnyampala, 2013b, pp. xi-xxvi), his work allows the reader to visualize exactly how these intricate rhyme chains are physically configured at a single glance for the *mtiririko* and *vidato* rhyme patterns[19]. The very existence of this sophisticated topological vocabulary demonstrates that the generative engine of the poetry is governed by strict syllabic principles, manifesting as a highly complex formal and spatial grid.

Before detailing some specific topological motifs, the relationship between the graphic grid and the poet's composition must be clarified. As established in Chapter 3 (section 3.3.3, part 4), the layout of the poem on the page does not preclude composition in the mind: the traditional poet (*mtunzi*), as Shariff reports, may measure the line and align the rhymes without pen or paper, drawing on the formulas and melodies of the tradition. What such a poet can hold in the mind, however, no protocol can tell us, and this study makes no claim about it. For later generations of formally educated poets, who compose directly on paper, the page itself serves as the primary generative

[19] For the purposes of this study, we exclude the *ngonjera* style invented by Mathias E. Mnyampala on the formal basis of the *shairi* meter, as well as the *msisitizo* meter, as they fall outside the scope of our classical corpus. However, it must be noted that these are innovations firmly rooted in tradition, continuing to operate strictly within the endogenous *vina na mizani* framework.

matrix of their syllabic architecture. In both cases, the structures are fixed on the page, and it is as they appear there that the following motifs are described.

Kupindulia (Class 15) (The Permutation / The Inversion)

Operating as a Class 15 nominal (infinitive/verbal noun), *kupindulia* literally translates to "changing direction," "turning," or "capsizing." It is derived from the Bantu verbal root *-pinda* (meaning "to turn" or "to bend"). Within the technical taxonomy of the *arudhi*, it designates specific metrical types (*aina*) within the broader poetic meters (*bahari*) characterized by an alteration in the orientation of the rhyme chains. *Kupindulia* describes a topological trajectory where the rhyme scheme alternates, inverts, or migrates across the lines of the stanza. Visually mapping these permutations on the graphic page often reveals intersecting paths resembling an "X" or an inverted "L" in romanized typescripts. The use of this term confirms a highly directional, spatial conception of the verse, completely detached from the linear progression of a temporal rhythm.

Mshuko (Class 3, pl. mishuko Class 4) (The Descent)

A nominal of Bantu origin derived from the verbal root *-shuka* (meaning "to descend" or "to go down"), *mshuko* translates literally as "a descent." Technically, it designates rhyme patterns that form continuous, parallel vertical columns extending downward across the stanza (*ubeti*). Whether these chains form double, triple, or quadruple vertical bars, the defining characteristic is their constant downward, vertical direction. This concept explicitly relies on the visual, graphic order of the written text: the rhyme chains are directed "downward" like visual pillars supporting the architecture of the stanza in the romanized typsecripts. The *mshuko* proves that the poet visualizes the poem as a descending spatial grid.

Kurejelea (Class 15) (The Recurrence)

Operating as a Class 15 nominal, *kurejelea* literally translates to "returning" or "recurring" (from the Arabic loan verb *-rejea*, "to return," fully assimilated into Bantu morphology). In the endogenous metrical framework, it describes the motif of structural recurrence, specifically when the terminal line—and its specific rhymes—repeats consistently throughout the entire poem as a refrain (Shariff, 1988, p. 50). By

rigorously observing this recurrent empirical data directly within the classical manuscripts, we were able to translate its mechanics into the formal models formalized in Chapter 8. It is vital to emphasize that this formalization operates strictly through an inductive process: the model is derived directly from the observable textual data, never the reverse.

Mtiririko (Class 3, pl. mitiririko Class 4) (The Continuous Flow)

Derived from the Bantu verbal root *-tiririka*, which means to "trickle," "flow," or "drip continuously" (like water), the *mtiririko* designates a highly demanding topological motif where the rhyme chains flow uninterrupted downwards across the entire length of the poem. In this spatial configuration, the poet maintains the exact same set of rhymes from the first stanza through to the final one.

To illustrate this continuous flow, Charles Mnyampala reproduces a masterful ten-stanza *shairi* composed by his father, Mathias Mnyampala, initially published in his *Diwani ya Mnyampala*, titled *Tungeni Mtiririko* ("[All of you] Compose a *Mtiririko*") (Mnyampala, 2013b, pp. xiii-xv). This configuration represents a more constrained variant of the *kupindulia* motif examined previously. In this poem, the spatial grid locks an "X" rhyme chain into a permanent vertical flow, reinforced by a completely immutable fourth line (*kibwagizo*) that repeats exactly in every stanza in Latin script. In the first three lines of each stanza, the medial rhyme on the first hemistich (*kina cha kati*) is fixed on 'bu', and the terminal rhyme on the second hemistich (*kina cha mwisho*) is fixed on 'ko'. On the fourth line, this topological trajectory permutes, completing the "X" motif: the 'ko' shifts to the middle and the 'bu' moves to the end (*Tungeni mtiriri**ko**, na silabi kuhasi**bu***).

In the tenth and final stanza, Mathias Mnyampala engages in a remarkable display of poetic metalanguage. He breaks the narrative frame to explicitly describe the exact spatial architecture and syllabic laws he has just composed:

> *Beti kumi kwa hesabu, hazina mchanganyiko,*
> *Nimemuweka Katibu, hapo kakaa kitako,*
> *Kati 'bu' imeratibu, mwisho 'ko' imekuweko,*
> *Tungeni mtiririko, na silabi kuhasibu.*
> (Mnyampala, 2013b, p. xv)

Ten stanzas by count, they have no mixture,
I have placed a Secretary, sitting right there,
In the middle, 'bu' is arranged; at the end, 'ko' is present,
[All of you] Compose a *mtiririko*, and count the syllables.

This final stanza acts as a direct, endogenous metapoetic manifesto. The poet visualizes the grid of the page explicitly, confirming the exact geometric mechanics of the rhymes. Furthermore, his use of the spatial words *kati* (in the middle) and *mwisho* (at the end) to describe the placement of 'bu' and 'ko' constitutes a brilliant metapoetic statement, perfectly mapping onto the endogenous technical terms for these structural positions: *kina cha kati* (the middle rhyme) and *kina cha mwisho* (the end rhyme). This vocabulary of place may also indicate how the poem was composed. A twentieth-century poet educated in Western-style schools, who describes his rhymes by where they sit in the line, appears to be working on the page, fully within graphic reason, and at some distance from the traditional poet described by Shariff, who composes in his head. This remains a hypothesis: the text shows how the poet describes his grid, not how he built it.

Of critical methodological significance is the repeated imperative refrain of the poem: *Tungeni mtiririko, na silabi kuhasibu* ([All of you] compose a *mtiririko*, and count the syllables). While proponents of the accentual school generally concede the presence of syllable counting in classical Swahili poetry, they systematically superimpose exogenous metrical feet (such as trochaic feet) onto the line to force a temporal rhythm. However, this primary evidence presents a fundamental contradiction to their premise: operating in the 20th century, a pioneering poet, heir to the classical tradition, explicitly commands a collective audience of aspiring poets to engage solely in the strict arithmetical counting of syllables (*silabi kuhasibu*). He offers no instruction to place beats, search for a metronomic pulse, or apply a stress pattern.

Kidato (Class 7, pl. vidato Class 8) (The Step / The Notch)

A purely Bantu nominal, *kidato* literally designates the "steps," "rungs," or "notches" carved into the trunks of palm trees to facilitate scaling. In Swahili metrics, *Mashairi ya Vidato* designates an incredibly complex generative architecture where stanzas structurally interlock like the rungs of a ladder. Instead of remaining isolated, each

stanza functions as a structural "step" within an ascending assembly. Through the sequential migration of the rhyme, a specific rhyming sequence from the preceding stanza shifts position to become the foundational rhyme of the subsequent stanza, chaining the entire poem together step by step (see QRG 4.3. for a formal overview, and 8.5.4. for a full description).

Precise historical and philological evidence confirms that the oldest known trace of this intricate *vidato* ("steps") creation style dates back to the 19th century, having been already perfectly executed by the classical Mombasa poet Muyaka bin Haji al-Ghassaniy (Hichens, 1940, pp. 92-93). The history of this interlocking motif shows that classical Swahili poets built their texts as architectures laid out on the page: whatever part of the work was done in the mind, it is on the page that the *vidato* can be seen and verified. Because the proponents of the exogenous accentual school are constrained by their search for temporal, rhythmic beats and acoustic stresses, they remain blind to this topological reality. The existence and early mastery of the *vidato* structure place the true generative engine of Swahili poetics far beyond the reach of these accentual theories, exposing accentuating metrics of classical Swahili poetry as an epistemological dead end.

4.3. The Paradox of Arudhi and Ubeti: Arabic Shells for a Swahili Metrical Core

The rigorous structural precision of the Swahili metrical grid is, paradoxically, rooted in a profound terminological dissonance. While the term *arudhi* (Class 9/10) is a manifest loan from the Arabic *ʿArūḍ* (عروض)—the "science of verse"—the borrowing remains purely nominal.

In the Swahili context, the term has been repurposed to describe a syllabic logic that formally negates the quantitative rules of its Arabic namesake. It functions as a lexical shell, emptied of its original quantitative substance to accommodate a Swahili phonological reality. Ultimately, the term *arudhi* was adopted solely as a prestigious label for a completely different generative grammar governed by syllabic constancy.

4.3.1. *Ubeti*: The Ultimate Lexical Shell

To grasp the mechanics of these lexical shells, we must examine the central unit of Swahili poetry: the *ubeti* (the stanza). Johnson's standard dictionary translates *ubeti* as "verse, stanza" and derives it from the Arabic *bayt* (بَيْتٌ) (Johnson, 1939); Sacleux glosses *beti* as "house" or "stanza/couplet," linking it to the Arabic *bayt*, "house; verse composed of two hemistichs" (Sacleux, 1939).

A single, direct borrowing from Arabic is only one possible history. The root, meaning "house," is Semitic: it is found in Amharic (ቤት / bēt), which also uses it for a unit of poetry (Leslau, 2004), and in Biblical Hebrew (בית / bayit, "house") (Sander & Trenel, 1859), where the poetic sense appears only later, in a medieval poetry modeled on Arabic. Comparative reconstruction does not trace the path a word has traveled; it measures the proximity between languages. What it shows here is that the root belongs to the Semitic branch of the Afro-Asiatic family, whose homeland Ehret (1995) places in Africa, near the Red Sea hills of Eritrea.

Other channels existed. The Persian world, bound up with Swahili history through the Shirazi identity of the city-states, down to Zanzibar's Afro-Shirazi Party, also builds its poetry on the *beit*, divided into hemistichs (*mesrā*) (Thiesen, 1982). Arabic influence itself came in distinct waves, such as the literary renewal brought by Hadhrami families in the 17th century (Vernet, 2005, pp. 162–164).

The word *ubeti* may therefore have reached the coast more than once, by different routes and through different contacts, across the Indian Ocean trade networks and a littoral of some 2,000 kilometers. This is a hypothesis, and it cannot be demonstrated from the word alone; but neither can a single borrowing. The burden of proof lies with whoever claims one source, such as a single borrowing from the Hadhrami families of the Lamu archipelago; and in any case, the history of the word says nothing about the meter it names.

4.3.2. The Architectural Metamorphosis

Regardless of its exact trajectory (or trajectories), the true nature of *ubeti* as a lexical shell becomes clear when we analyze its structural metamorphosis.

In both the Arabic and Persian classical traditions, the *bayt* or *beit* designates a single line of poetry. As Georges Bohas and Bruno Paoli synthesize, the classical Arabic metric system does not operate at the

stanza level; its primary unit is the single line (*bayt*), which is strictly bipartite, divided into a first hemistich (*al-ṣadr*, "the chest") and a second hemistich (*al-ʿajuz,* "the behind") (Bohas & Paoli, 1993). These hemistichs are built upon metrical feet (*tafʿīla*), each combining a disyllabic core called a *watid* ("peg") with one or two elements called *sabab* ("cord"), a *sabab* being either one long syllable or two short ones. The key differentiation between these feet relies entirely on the distinction between short and long syllables. Classical Persian metrics operates on the same foundational logic: the *beit* is divided into two *mesrā*, each containing metrical feet (*rokn*, pl. *arkān*) strictly defined by the opposition of short and long syllables (Thiesen, 1982).

When Swahili poets adopted this terminology, they executed a drastic architectural expansion (see Chapter 8 for a detailed formalization of this Swahili architecture, or the Quick Reference Guide for a formal overview). In the Swahili metrical grid, an *ubeti* is almost never a single line. It is a complex, fixed, multi-line stanza (for example, spanning four, five, or six lines). Furthermore, the internal division of the Swahili line (*mshororo*) into segments (*vipande*) breaks with the strict bipartite mirror of the Arabo-Persian hemistich. While the standard 16-syllable shairi bisects its lines into two equal *vipande* (eight and eight syllables), other meters and sub-meters do not. An *utenzi* line is undivided. A *wimbo* line can be divided by internal rhyme chains into two to four sections (*vipande*) of unequal syllabic lengths. The number of parts per line in Swahili is functionally variable and meter-dependent.

Thus, the overarching Swahili stanza is structured by arithmetical syllabic constraints (*mizani*) and interwoven rhyme chains (*vina*) that dictate the lines and their internal sections. The architectural complexity of the Swahili *ubeti*, repeating a specific metrical motif across multiple lines to form a stanzaic block, far exceeds the simple linear bipartition of its Semitic or Persian namesakes. The Swahili poet took the name of a foreign structural brick and repurposed it to designate an entire endogenous building.

4.3.3. The Illusion of a Total Borrowing

This terminological adoption creates a dangerous epistemological trap for exogenous scholars: the illusion of a total borrowing.

Because the Swahili system utilizes foreign names like *ubeti* and *arudhi*, a specific, albeit historically tenacious, faction of exogenous analysts have leapt to the conclusion that Swahili must have also

imported the foreign phonological engine that powered the original *bayt*. They falsely assume a total cultural transfer that would go so far as to redefine the fundamental phonology of the Swahili language, projecting classical Greek or Arabic metrics onto Swahili texts.

It is essential to nuance this: not all Western scholars fell into this epistemological trap. Scholars such as Lyndon Harries, Jan Knappert, and others recognized the distinct nature of Swahili poetry without forcing it into an exogenous metrical mold. However, the quantitative and accentual schools have proven remarkably persistent over the last century, sometimes resorting to intense intellectual pressure and distorting linguistic reality to ensure their own survival.

This total borrowing is structurally and linguistically impossible. The Swahili poet utilizes the prestigious lexical shell of the *bayt*, but strips it of its quantitative substance. As linguists have established, Swahili has no phonological opposition between long and short syllables (Sacleux, 1909, pp. 6-10; Ashton, 1962, pp. 3-5; Polomé, 1967, pp. 36-59). The only vocalic lengthening in Swahili is a localized phonetic phenomenon triggered by the penultimate stress accent, lacking any distinguishing phonological value. Without short and long syllables, the entire polysyllabic foot system (*tafʿīla / rokn*) of classical Arabic and Persian metrics collapses. It is impossible for Swahili poetry to utilize the same metrical engine as the Arabs and Persians at the elementary foot level.

Furthermore, this strict syllabic logic does not call for accentual or syllabo-tonic metrics (stress-based feet) either. Recent scholarship, such as Emiliano Minerba's *Meters in Motion* (Minerba, 2025), still attempts to impose an accentual framework on classical Swahili poetry. Minerba explores the Arabic metrical foot (*tafʿīla*) through Gotthold Weil's theory, which proposes an "accentual interpretation of the Halilian system" by assigning an ictus (dynamic stress beat) to the core of the metric foot. Building on this pursuit of stress-based rhythms, Minerba categorizes the development of the classical Swahili meters (*utenzi*, *shairi*, and *wimbo*) as a "trochaic innovation" (see section 2.4 for a detailed analysis of Minerba's framework).

However, this syllabo-tonic approach is yet another epistemological illusion. The Swahili identity is characterized by deep historical *métissage*, making the adoption of Arabic terminology perfectly natural for an Indian Ocean civilization. Yet this influence remains purely meta-linguistic. The vocabulary used to talk about the poetry is borrowed, but the deep architecture itself is generated by syllabic principles conforming to the phonology of Swahili, a Bantu language.

The generative grammar inside the *ubeti* remains strictly Swahili: it is built not on the duration of the voice (quantitative), nor on the dynamic rhythmic stress of a metrical foot (accentual), but exclusively on the arithmetic count of the syllables (*mizani*) and the structural logic of the rhyme chains (*vina*), which actively carve out the lines and the hemistichs.

This deep semantic duality is brilliantly encapsulated in the Swahili terminology itself. The word for rhyme, *kina* (plural *vina*), presents a remarkable ambiguity. In its Arabic etymology, it refers to the rhyme. But in its Bantu etymology, *kina* signifies "deep water" or "depth." This semantic co-variation mirrors the word *bahari*, an Arabo-Persian loanword meaning "ocean" or "sea," which Swahili poets use to designate the poetic meter. However, unlike *ubeti* (*bayt*), whose root, meaning "house," is shared by the Semitic languages, while the history of its poetic sense remains uncertain, the dual meaning of *bahari* is the result of direct cultural diffusion. In classical Arabic, the term *baḥr* (بحر) designates both the sea and the metrical pattern—a specific intellectual metaphor established by 8th-century Arab prosodists to evoke the vastness of poetic creation, which subsequently spread throughout the Islamic world. Thus, a Swahili phrase like *bahari ina kina* contains a stunning double entendre: it translates both to "the meter has a rhyme" and "the ocean has depth." Classical Swahili poetry is much like this phrase: Arabo-Persian on the surface of its meta-linguistic descriptions, but Swahili in its metrical depths.

4.4. The Epistemological Arbiter: *Guni* (The Defective Verse)

The most devastating lexical proof against the accentual school is the concept of *guni*.

Within accentual Western paradigms, deviations from strict meter are often celebrated as rhythmic variation, artistic license, or syncopation. In the classical Swahili matrix, any failure to uphold the absolute syllabic count (*mizani*) or the rhyme scheme (*vina*) is not an alternative artistic path; it is a clinical defect.

Shaaban Robert (1972) defines the boundaries of *guni* with uncompromising clarity: "*Mashairi ya guni hujulikana kwa dosari ya mizani na ila katika vina*" ("Poems of guni are known by the defect of measure and the flaw in rhymes").

In the entirety of classical Swahili poetic scholarship, a line is categorized as *guni* (flawed or defective) only when the absolute syllable count (*mizani*) fails or when the terminal homophony (*vina*) is broken.

There is no term in the Swahili language to describe a "rhythmic failure." A verse is never declared *guni* because a supposed "rhythmic beat" was missed, or because a performer failed to syncopate an accentual foot. The concept of rhythmic failure does not exist in the formal evaluation of the poem, proving that the generative engine is exclusively arithmetic and syllabic.

4.5. The Division of Labor: Composition vs. Performance

4.5.1. The Architects of the Compositions

Mtunzi (Class 1, pl. watunzi Class 2) (The Composer / The Poet-Architect)

Derived from the Bantu verbal root *-tunga* (meaning "to compose," "to arrange," or "to string together like pearls," which directly yields the nominal *tungo* analyzed above). In its broadest, common, and contemporary sense, a *mtunzi* is simply a composer, a creator, or an author. This modern usage is highly inclusive, encompassing anyone who produces literary or musical works, including contemporary poets who champion unregulated free verse (*mashairi huru*) and actively reject traditional metrical constraints. However, within the strict, classical taxonomy of *arudhi* (metrics) that forms the core of this study, the *mtunzi* is specifically the poet-architect, the intellectual creator and composer of the verse. In this classical context, while serving the poetic message, their primary structural task is the exact arrangement of the syllabic measure (*mizani*) and the precise anchoring of the rhyme chains (*vina*). The classical *mtunzi* operates exclusively in the realm of verbal artifact and structural alignment; they construct the invariant foundation of the poem. Within the classical Swahili tradition, the individual who composes this architectural blueprint does not necessarily perform it. The classical *mtunzi* provides the immutable composition, establishing a rigid, predictable foundation that is subsequently handed over to the performer.

Mshairi (Class 1, pl. washairi Class 2) (The Poet)

Derived from the Arabic root *shi'r* (poetry) and fully assimilated into Bantu morphology, the term *mshairi* translates basically to "poet". Much like *mtunzi*, its everyday, modern usage serves as a universal umbrella term for anyone composing poetry, regardless of their adherence to classical rules. Today, a *mshairi* can perfectly well be a practitioner of free verse who operates entirely outside the boundaries of syllable counting and rhyming. Nonetheless, within the strict formal framework of classical *arudhi*, *mshairi* functions as a synonym for the classical *mtunzi*. It designates the practitioner as the creator of the poem, responsible for both its semantic content and its formal architecture. The classical *mshairi* conceives and structures the work, ensuring the precise organization of its syllabic measure (*mizani*), rhyme system (*vina*), and poetic expression.

As Lyndon Harries explained: "[...] the poet almost invariably employed a copyist or a narrator who memorized the poem and later wrote it down" (Harries, 1962, p. 5). The *washairi* are the architects of the poem, whose task ends with the completion of its verbal composition. According to the theory advanced by Swahili scholar and poet Ibrahim Noor Shariff (1988), for traditional poets at least, this architectural process operated strictly within the mental realm. In this view, the poet composed the intricate structure of the lines entirely in the mind, with the physical act of writing intervening only after the cognitive composition was complete. This subsequent transcription into manuscript form served precise functional purposes: to preserve the work for memory and posterity, to dispatch it across distances to correspondents, or to provide a definitive script for recitation. Shariff acknowledges, however, that poets influenced by Western education composed directly on the page. Yet, whether the manuscript serves as a subsequent transcription or the primary medium of creation, both methods converge on the production of a fixed text. Consequently, its vocal realization, grounded in that fixed written composition, belongs entirely to the performer (*mghani*).

Exogenous scholars frequently err by romantically assuming the *mshairi* is a bard or a wandering minstrel[20]; in the endogenous classical reality, the *mshairi* is an intellectual composer strictly bound to the uncompromising laws of *vina na mizani*.

[20] e.g., "[Alice] Werner compared *tendi* poems to the 'loose rhymes' of the Welsh bards, used by them for fluent narrative of a less elevated kind. [sic]" (Harries, 1962, p. 25)

Malenga (Class 1, pl. malenga Class 2) (The Master Poet / The Virtuoso)

A purely Bantu nominal designating an expert poet, a supreme virtuoso of the *vina na mizani* system. A *malenga* is a *mshairi* who has achieved absolute mastery over the most complex metrical configurations (such as the interlocking *vidato* or intricate asymmetrical hemistiches). The title is one of immense intellectual prestige, acknowledging a composer whose structural command of the poetic form is flawless. Like the classical *mtunzi* and the *mshairi*, the *malenga* belongs strictly to the domain of composition and textuality. They are the master architects of the classical tradition, distinct and entirely separate from the performers (*waghani*) who will eventually vocalize their blueprints.

4.5.2. The Performative Realm (The Execution)

Mghani (Class 1, pl. waghani Class 2) (The Performer / The Singer)

Derived from the verbal root *-ghani* (meaning "to recite," "to chant," or "to sing poetry"). The *mghani* is the professional performer or singer who executes the vocal and acoustic realization of the poem. While the *mtunzi* conceives the poem's meaning and constructs its invariant metrical framework, the *mghani* provides the melodic and aesthetic overlay. The *Kamusi ya Kiswahili Sanifu* (2004) published by Oxford University Press defines the *mghani* as: "*Mtu anayeimba mashairi*" ("A person who sings poems"). Their artistic sovereignty lies precisely in their reliance on the rigid, predictable syllabic measure fixed on the written page. This invariant foundation grants them absolute freedom to stretch vowels, introduce syncopation, and superimpose complex vocal melodies without ever risking the structural collapse of the poem.

Mahadhi (Class 6) (The Melody / The Performative Rhythm)

A nominal designating the performative melody, the tune, or the musical rhythm actively applied during the oral execution of the text. *Mahadhi* is strictly a melodic overlay belonging to the singer, completely distinct from the invariant metrical structure crafted by the *mtunzi*. While the architectural blueprint (*vina na mizani*) is an

absolute syllabic law, the *mahadhi* is highly variable; a single classical poem with a perfectly stable syllabic foundation can be sung to dozens of different traditional melodies. While further ethnolinguistic and musicological research is undoubtedly required to fully map the complex variations and historical origins of these specific traditional tunes, for the rigorous purpose of formal metrics, the fundamental epistemological requirement is simply to maintain the absolute separation between the invariant structural composition and its fluid musical performance. The *mahadhi* belongs to the domain of sound, not to the geometry of the page.

4.6. The Endogenous Accentual Lexical Void and Academic Projections

Because classical Swahili poetics operates as a purely formal, structural, and spatial architecture, its technical lexicon serves as a direct, empirical mirror of its generative cognitive framework. In the study of specialized terminology, what a lexicon includes is just as epistemologically telling as what it systematically excludes.

Swahili poets forged and assimilated a highly granular vocabulary to map every coordinate of their verse. This spatial dimension operates as an intricate interplay between internal mental architecture and external graphic reality. Yet, as established earlier, our methodological access to this architecture is strictly confined to the written artifact. Lacking any empirical entry into the poet's mind during composition, we must rely exclusively on the page.[21]

To navigate this dual mental and graphic matrix, the endogenous vocabulary deploys surgical precision, mapping the architecture of the verse across highly specific conceptual domains.

First, it encompasses a rigorous counting paradigm dedicated to the arithmetic of the poem, featuring explicit concepts for exact syllable counting, exact measurement, metric equilibrium, and the specific

[21] While some poets could engineer their verses entirely within a cognitive grid prior to any physical transcription or dictation, the existence of highly complex topological forms raises a profound theoretical question. For such intricate architectures, the technology of writing itself—and the cognitive spatialization it affords, precisely echoing Jack Goody's concept of "graphic reason"—likely played an active role in structuring the poem's spatiality, fundamentally extending the poet's mental capacity. However, this theoretical question remains open.

enumeration of lines per stanza. Second, it provides a comprehensive terminology for the poem's formal anatomy, mapping the composition at every scale. It defines the macro-level of meters, zooming in on stanzaic units, individual lines, and their internal subdivisions into hemistichs and smaller sections. Furthermore, it explicitly names the structural "currents"—the specific syllabic spans that dictate the exact length of these line segments. Third, the lexicon strictly categorizes all metric locks, boundaries, and trajectories, providing dedicated terminology for rhymes, strict structural syllabic measures utilized in the absence of rhyme, and specific resting points or caesuras. Finally, it possesses an entire vocabulary dedicated to topological motifs and spatial architecture, explicitly naming the geometric mechanics of rhyme schemes—such as interlocking structures, returning cyclical motifs, spatial reversals, descending vertical flows, and stepped, staircase-like structural formations.

Yet, despite possessing this exhaustively detailed, multidimensional structural lexicon to formally construct and spatially navigate the poem, there is no word in Swahili to designate a metrical foot or a rhythmic beat.

This conceptual absence is corroborated by a striking and highly specific lexical void. While classical Swahili heavily borrowed foundational poetic terminology from Arabic and Persian to describe its overarching structural framework—adopting concepts for the metrics *(arudhi)*, the stanza *(ubeti)*, and the rhyme *(kina)*—it completely ignored the Arabic term for the metrical foot, *tafʿīla* (تَفْعِيلَة). This omission is not a matter of linguistic incompatibility. Phonologically, this word could have been assimilated into the Bantu open-syllable (consonant-vowel) system, likely adapting into a smooth three-syllable form such as *tafila*. The fact that Swahili poets integrated a vast exogenous vocabulary for their prosody but entirely bypassed the specific terminology for metrical feet speaks for itself. It demonstrates that superimposed temporal rhythms and accentual feet were not simply "lost in translation" or accidentally forgotten; they were never part of the endogenous cognitive framework to begin with. Vocabulary follows operational necessity, and the Swahili poet simply had no use for a tool designed to measure a rhythm that did not exist in their lines.

Consequently, there is no term in the Swahili language to describe a "rhythmic failure." A classical verse is never declared *guni* (defective) because a supposed "rhythmic beat" was missed, or because a performer failed to syncopate an accentual foot. The

concept of rhythmic failure does not exist in the formal, endogenous evaluation of the poem, proving that the generative engine is exclusively arithmetic and syllabic. Swahili poets systematically excluded the concept of the foot from their system because their internal weighing relies on the static syllabic measure of the line, not the dynamic stepping of a metrical foot. There is no equivalent to the Greek πούς (*pous*, literally meaning "foot," the basic rhythmic unit of quantitative measurement) because classical Swahili poetry does not "walk" to a beat; observable on the page, it balances syllabic measures on a spatial grid.

Confronted with this endogenous lexical void, some exogenous scholars have felt compelled to import foreign terminology to force a temporal rhythm onto a syllabic architecture. The historical application of Greek prosodic models—such as Büttner (1894) defining the meter as iambic or trochaic, Steere (1870) as anapaestic, or Allen (1971) as ionic—demonstrates the complete structural limitation and fundamental incompatibility of foreign metrical theories. When modern scholars like Clarissa Vierke (2011, pp. 39, 54) state that the Swahili line mimics the Aristotelian paean—παιάν (a term used in prosody to denote a four-syllable foot)—she inadvertently exposes the methodological collapse of her own paradigm. By attempting to force an accentual pulse into a system governed solely by syllabic principles, these analyses invent a structural problem that classical Swahili poetics never possessed, only to offer an imported solution that the poetry never needed.

To supplement the sovereign Swahili metrical architecture with metrical feet imported from Greek terminology (such as trochees) or to force it into an Arabic mold—artificially retrofitted with Germanic accentual mechanics by substituting the unworkable Arabic vocalic length with rhythmic stress beats labeled as a trochaic innovation—is not a philological clarification; it is an academic projection. The *vina na mizani* framework requires no imported parts. Rooted entirely in its endogenous syllabic principles, it is structurally complete, formally perfect, and autonomous.

Chapter 5. The Empirical Reality of *Vina na Mizani*: Voices of the Master Poets of the Swahili Language

The fundamental irony of the prolonged debate over classical Swahili metrics is that the system has never been a mystery awaiting decipherment by Western philologists. It is, and has always been, a highly formalized, self-documented discipline. Swahili-language intellectuals, poets, and lexicographers have consistently described their own art form using a precise, formal vocabulary explicitly designed to codify the boundaries and syllabic exactness of their verses. The persistent failure of exogenous theorists to acknowledge this endogenous theoretical framework is not a mere methodological oversight; it is a characterized epistemological erasure.

To accurately model the classical Swahili metrical system, we must abandon the imported tools of ancient Greek or syllabo-tonic Arabic accentual prosody. Instead, we must listen directly to the creators and scholars themselves. When we allow the master poets to speak, a massive, intergenerational, and highly detailed consensus emerges—one that spans decades of scholarship and defines the classical Swahili poetry exclusively by its uncompromising formal boundaries: the absolute syllable count (*mizani*) and the rhyming homophony (*vina*). By systematically analyzing the original Swahili texts of Sheikh Kaluta Amri Abedi, Shaaban Robert, Ibrahim Noor Shariff, Ahmed Sheikh Nabhany, Hashil Seif Hashil, and Mathias E. Mnyampala, we restore the sovereign endogenous voices that the accentual school has so long ignored.

One objection must be met before the evidence is presented, because it determines what that evidence can establish. It may be said that the silence of the poets and intellectuals who have reflected on the tradition on the subject of accent proves nothing, since Swahili stress is automatic: it falls on the penultimate syllable of every phonological word, no composer needs to be taught this, and writings on the craft have no reason to state a rule for what the language does by itself. The objection would be decisive if the accentual school claimed only that stress exists. It does not. It claims that the eight-

syllable line contains exactly two beats, in preferred positions, and that this configuration is a cognitive prototype toward which the composer works. That is not an automatic property of the language: it is a selection among the lines the language makes available, and therefore a constraint on the choice of words. A constraint on the choice of words is precisely the kind of rule that these practitioners and scholars spend their texts stating, in a tradition that codifies with great precision how many syllables a line must have, where the rhyme must fall, and what happens when either fails. The silence bears, then, not on the existence of stress, which no one disputes, but on the alleged constraint, which is exactly where the accentual hypothesis has content.

5.1. Sheikh Kaluta Amri Abedi and the Absolute Syllabic Foundation (1954)

The formalization of modern Swahili metrical theory was inaugurated by the distinguished poet, scholar, and future mayor of Dar es Salaam, Sheikh Kaluta Amri Abedi. His seminal treatise, *Sheria za Kutunga Mashairi na Diwani ya Amri* [The Rules of Poetry Composition and the Anthology of Amri], published in Tanganyika in 1954, established the definitive theoretical foundation for generations of poets across East Africa. Abedi did not invent the system; rather, he was the first to systematically codify a centuries-old practice, translating the esoteric knowledge of the coastal elites into an accessible, formal metrical framework for the broader Swahili-speaking public.

Abedi's entire methodological approach is taxonomic and relies exclusively on the exact counting of formal units. He categorizes the classical meters entirely based on the fixed number of lines (which he terms *mistari*, singular *mstari*) per stanza (*ubeti*, plural *beti*). In his formal classification, he identifies the *Tathlitha* (stanzas composed of exactly three lines), the *Takhmisa* (stanzas of five lines), and the *Tenzi* and *Mashairi* (which both utilize four-line stanzas i.e. *Tarbia* but are strictly differentiated by their exact syllable counts and rhyme chains). To describe the internal division of the line into hemistiches, he uses the term *upande* (plural *pande* or *vipande*, literally meaning "the sides" or "parts"). At no point in this exhaustive formal taxonomy does he mention a temporal pulse, a rhythmic beat, or an accentual foot.

5.1.1. The Technical Definition of the *Utenzi* Parameters

To establish this taxonomy, Abedi recognizes the generative fluidity of the tradition, explicitly linking the diversity of poetic forms to the autonomous intellect and craftsmanship of the composer:

> Poetries are of many kinds. We cannot say immediately how many kinds there are because it is a thing that depends on the skill of a person's intelligence. (Abedi, 1954, p. 16)[22]

Abedi then provides a highly specific technical definition for the *utenzi* meter, anchoring it firmly in a grid-based, syllabic reality:

> The word, *Utenzi*, here its meaning is a poetry that has fewer than twelve measures, which does not have median rhymes in the lines, but each stanza has rhymes of a single kind in its lines if not for the final line in each stanza, which develops with a rhyme of another kind that is called *bahari*. (Abedi, 1954, p. 29)[23]

From this precise description, we can extract the strict defining traits of the *utenzi* meter according to Abedi: Short graphic lines (*mistari*) structure the stanza (*ubeti*); these lines have a number of syllables less than *(chache kuliko)* twelve. The graphic lines do not have median rhymes (*vina vya kati*). Two different rhyme chains are present in the stanza (*ubeti*), one of which has as its specific place of realization the last graphic line of the stanza (*ubeti*).

---

[22] Swahili original : "Mashairi yako namna nyingi. Hatuwezi hivi sasa kusema ni ya namna ngapi kwa vile ni kitu kinachotegemea usanifu wa akili ya mtu."

[23] Swahili original : "Neno, Utenzi, hapa maana yake ni ushairi wenye mizani chache kuliko kumi na mbili, ambao hauna vina vya kati katika mistari, bali kila ubeti una vina vya namna moja katika mistari yake isipokuwa mstari wa mwisho katika kila ubeti, ambao huendelea na kina cha namna moja kinachoitwa bahari."

5.1.2. The Ostensive Definition: Regulated Metrical Variations

To substantiate this theoretical framework, Abedi provides an ostensive definition, demonstrating the system in action through concrete, empirical instances. These examples prove that the syllabic measure (*mizani*) is a structural variable governed by clear rules. He documents traditional variations operating at 6, 8, and 11 syllables per graphic line, designated in Swahili as *Tenzi za mizani 6* (*Tenzi* of 6 syllables), *Tenzi za mizani 8* (*Tenzi* of 8 syllables), and *Tenzi za mizani 11* (*Tenzi* of 11 syllables).

Example at six syllables, one stanza (*ubeti*):

> **L**1: We-ngi hu-ji-vu-**na,** [6]
> Many boast,
> L2: Hu-zu-sha fi-ti-**na,** [6]
> Arising quarrels,
> L3: Na nyi-ngi hi-ya-**na ,**[6]
> And much deceit,
> L4: Zi-ki-wa-zu-a-zu-**a.** [6]
> Which slander them.
> (Abedi, 1954, p. 32)

While the stanzaic rhyme scheme **na** is purely local, internal and variable from one stanza to another, the final rhyme **a** remains constant across every stanza throughout the poem.

Example at eight syllables, one stanza (*ubeti*):

> L1: Ha-ya ma-mbo ya ku-tu-**nga**, [8]
> The matters of composition,
> L2: Sha-ri-ti u-we ma-le-**nga**, [8]
> You must be a master poet,
> L3: A-u ma-ne-no ya Ma-**nga**, [8]
> Or the words of Manga [the North],
> L4: U-ya-ju-e u-dhu-i-**ha**. [8]
> That you know them clearly.
> (Abedi, 1954, p. 30)

While the stanzaic rhyme scheme **nga** is purely local, internal and variable from one stanza to another, the final rhyme **ha** remains constant across every stanza throughout the poem.

Example at eleven syllables, one stanza (*ubeti*) — *Utenzi wa Inkishafi*:

> L1: Ba-si Mu-o-ngo-zi Mo-la Ra-hi-ma, [11]
> Well, the Guide the Lord the Merciful,
> L2: Hu-tu-ma wa-tu-mwi wa-li-o we-ma, [11]
> Sends messengers who are good,
> L3: Wa-ka-i-te wa-tu kwe-nye sa-la-ma, [11]
> That they then call the people into peace,
> L4: Wa-u-po-ne mo-to u-u-ngu-za-o. [11]
> That they heal from the fire that burns.
> (Abedi, 1954, p. 31)

While the stanzaic rhyme scheme **ma** is purely local, internal and variable from one stanza to another, the final rhyme **o** remains constant across every stanza throughout the poem.

Notwithstanding the absolute syllable count of each line (whether 6, 8, or 11), the syllabic rhyme chain binding lines 1 to 3 operates with an internal scope within each individual stanza, and its phonetic nature can vary from one stanza to another. Conversely, the final rhyme of line 4 describes a continuous rhyme chain that remains strictly identical across every single stanza of an *utenzi* poem; this omni-strophe continuity constitutes the very elegance of the meter, acting like the thread of a necklace where the strophes are woven together like pearls.

It is worth noting that this definition is not merely compatible with the formal model developed in Chapter 8: it generates it. Abedi's upper bound of twelve *mizani*, together with the three configurations he documents, yields directly the syllabic parameter $N \in \{6, 8, 11\}$ of the *utenzi* matrix given in QRG.2 and section 8.3. The matrix is a formalization of an endogenous definition published in 1954, not a modern reconstruction imposed on the corpus.

5.1.3. The Syllable as the Sole Metric Unit

Throughout these exhaustive technical definitions and empirical demonstrations, we must underline the complete absence of any notion of an accentual foot or of vocalic quantity. The system is completely syllabic. When defining the internal generative mechanics of the Swahili verse, Abedi explicitly identifies the syllable (*silabi*) as the sole metric unit that locks the formal grid into place. While the *utenzi* lacks internal rhymes, Abedi demonstrates this absolute

syllabic reliance when defining the *shairi* meter, a genre defined precisely by the presence of these median rhymes (*vina vya kati*). When describing the mechanics of the rhyme (*kina*, pl. *vina*) across the lines of a *shairi*, he writes with uncompromising formal clarity:

> [...] usually the entire poem has two syllables – in the middle and at the end – which occur again and again even to the point of annoying the ear. (Abedi, 1954, p. 20)[24]

This definition is epistemologically devastating for the exogenous accentual school. Abedi specifies that the formal layout of the *shairi* is secured by the exact repetition of specific syllables at precise spatial boundaries: the caesura in the middle of the line and the terminal rhyme at its end. The foundation of the verse is defined by strict syllabic recurrence by this master poet, not by temporal or accentual feet. There is no mention of an underlying rhythm that compensates for missing syllables; if the exact syllable is missing, the formal contract is simply broken. The repetition is so absolute and phonetically demanding that Abedi notes it can almost "annoy the ear"—a direct consequence of severe formal constraints, not of a loose, melodic, stress-timed rhythm.

5.1.4. An Autonomous Intellectual Weapon

Finally, Abedi understood that this rigorous, endogenous formal system was not merely an aesthetic curiosity; it was an autonomous intellectual weapon. The formal perfection of the Swahili language and its poetry provided a sovereign, unifying force that proved African intellectual capability and defied colonial hegemony. In his introduction to Mnyampala's *Waadhi wa Ushairi*, Abedi elevates the Swahili language to a gift of God, an unifying factor and an anti-colonial force:

> Every good praise belongs to Almighty God who granted that Swahili be spoken on the Mainland and the Coast. By it, he united the original inhabitants and the settlers who differed in tribes, color, and religion. He enabled the leaders to use Swahili to open people's minds, and they were illuminated by the dawn of realization followed by a morning shining with a multitude of rays

[24] Swahili original : "[...] huwa shairi zima linakuwa na silabi mbili – kati na mwisho – zinazotokea tena na tena hata kwa kiasi cha kuliudhi sikio."

> and perfect light. By it, the weak gained strength, and the resilience to fight the authoritarian Colonial Government. (Mnyampala, 2011, p. 103)[25]

Abedi presents the formal Swahili language, perfected in its exacting poetic meters, as an autonomous African engine, strong enough to unite a diverse nation and to sustain the struggle against colonial authority. This does not, by itself, settle anything about the structure of the verse: a meter is not made syllabic by the political use to which its language is put. But it makes the operation of the accentual school something more than a technical error. What is being superimposed on this system is not a neutral instrument. It is an analytical apparatus formed in the philology of the colonial period, applied to a formal architecture that its own codifiers had defined, in the same decades, as the guarantee of their intellectual autonomy. The two claims of sovereignty are not compatible, and the discomfort this produces is not an argument, but it is a fact about the situation that no account of this scholarship can honestly leave out.

5.2. Shaaban Robert and the Ontological Necessity of the Syllabic Matrix

If Sheikh Kaluta Amri Abedi provided the technical grammar of classical Swahili poetics, it is Shaaban Robert who elevated this formal framework to the level of an ontological necessity. Operating in profound continuity with the classical heritage—even though he is considered by some strictly as the father of *modern* Swahili literature—the poet laureate of East Africa does not treat the endogenous rules of *vina* (rhymes) and *mizani* (syllabic measure) as mere aesthetic options; rather, he posits them as the absolute structural bedrock of all literary education and human expression.

Before even defining the metrical structures, Shaaban Robert establishes the monumental gravity of the poetic form with a striking

[25] Swahili original : "Kila sifa njema ni ya Mwenyezi Mungu Aliyejalia Kiswahili kisemwe Bara na Pwani. Kwacho akaunganisha wakazi wa jadi na walowezi waliohitilafiana makabila na rangi na dini. Akawawezesha viongozi kutumia Kiswahili kiwafungue watu akili, wakapambazukiwa na alfajiri ya utambuzi iliyofuatwa na asubuhi inawirishayo kwa wingi wa miali na mwangaza kamili. Kwacho wanyonge wakapata nguvu, na ukakamavu kupambana na Serikali ya Kikoloni yenye mabavu."

geographic metaphor in his introduction of *Sanaa ya Ushairi* (Robert,1972), his book dedicated to the Art of Poetry:

> Poetry is one kind of art that is essential in education. Education without poetry is identical to Egypt without the Nile. Moreover, poetry is an innate nature that a person is born with. (Robert, 1972, p. xii)[26]

By equating the art of poetry to the Nile River, Shaaban Robert asserts that the intellectual lifeblood of the Swahili world flows naturally through these verses, establishing the poetic form not merely as a cultural artifact, but as a fundamental ontological necessity of human nature itself. Having established this supreme dignity, he proceeds to categorize the tradition. Just like Abedi, his entire taxonomy of poetic forms is articulated around the rigorous, formal presence or absence of exact syllabic counts and phonetic homophonies, completely ignoring any exogenous notion of accentual rhythm.

5.2.1. *Mashairi ya Mizani*: The Absolute Measure

Shaaban Robert's definition of meter is syllabic. Under the heading *Mashairi ya mizani* (Poems of measure), he defines the metric foundation not as a sequence of musical beats, but as a strict physical measurement of linguistic units: "These are poems having a measure or measures in the pronunciations of characters or syllables." (Robert, 1972, p. xii)[27]

He leaves no room for rhythmic approximation or the stretching of vowels to fill a theoretical temporal void. For Shaaban Robert, poetic excellence relies on structural perfection, which he explicitly identifies as *mizani kamili* (perfect/complete measure). The architecture of the stanza is defined by this numerical absolute.

[26] Swahili original: "Ushairi ni namna moja ya sanaa muhimu katika masomo. Masomo bila ya ushairi ni sawa na Misri bila ya Nili. Zaidi ya hayo ushairi ni maumbile ambayo mtu huzaliwa nayo."

[27] Swahili original: "Haya ni mashairi yenye kipimo au vipimo katika matamko ya herufi au silabi."

> Some of these poems usually have stanzas of a single kind of measure throughout the entire work. A measure of a single kind is identical to saying a perfect measure. (Robert, 1972, p. xii)[28]

To demonstrate this perfect structural modularity, he lists specific numerical variants (Robert, 1972, p. xii): an exact symmetry of "eight by eight measures" (*mizani nane kwa nane*) or "sixteen by sixteen measures" (*mizani kumi na sita kwa kumi na sita*), as well as complex asymmetrical but strictly regulated balances like "seven by eight measures" (*mizani saba kwa nane*) or "eight by six measures" (*mizani nane kwa sita*). The metric engine of Shaaban Robert is purely arithmetical and syllabic.

5.2.2. The Architecture of *Vina* and the *Habari* of the *Utenzi*

Equally foundational to the syllable is the phonetic lock of the rhyme (*vina*). Shaaban Robert defines *Mashairi ya vina* (Poems of rhymes) through the precise spatial alignment of sound at specific boundaries within the grid:

> These poems have a harmony of sound in the middle of and at the end of every line of the stanza. (Robert, 1972, p. xiii)[29,30]

When addressing the *utenzi* meter specifically, Shaaban Robert introduces a fascinating terminological precision. While modern analysis might use abstract variables or Abedi's term *bahari* to

---

[28] Swahili original: "Baadhi ya mashairi haya huwa na beti za mizani namna moja katika kazi nzima. Mizani namna moja ni sawa na kusema mizani kamili."

[29] Swahili original: "Mashairi haya yana mlingano wa sauti kati ya na mwishoni mwa kila mstari wa ubeti."

[30] Literally speaking, Shaaban Robert's definition of "poems of rhymes" describes a metrical architecture built upon both median and final rhymes. However, immediately following this definition, he provides an example that completely lacks median rhymes, explicitly introducing it as "*Mfano wa shairi la vina mwishoni* [An example of a poem with rhymes at the ends]". Rather than viewing this omission as a theoretical contradiction, this partial exemplification confirms that his initial definition functions as an inclusive, generative matrix. As a master of the language, he implicitly treats the median boundary not as a rigid obligation, but as an available metrical parameter. Thus, his overarching definition establishes the maximum architectural potential of the verse: the final rhyme acts as the absolute structural constant, while the median rhyme operates as an optional variable that can be activated or omitted depending on the specific composition being executed.

describe the constant, macro-structural terminal rhyme that spans the entire poem, Shaaban Robert conceptualizes this continuous rhyme chain as the *habari*:

> The *utenzi* is a poem having a *habari* in the final line of every stanza. The *habari* in the art of poetry is a kind of rhyme that continuously occurs in every final line of the *utenzi*. (Robert, 1972, p . xiii)[31]

It is important to note that Shaaban Robert strips the word *habari* of its standard semantic meaning ("news" or "information") to employ it in a strictly technical capacity. In his specific metrical vocabulary, the term is devoid of any narrative weight; it simply designates the omnistrophic terminal rhyme, the continuous phonetic constant that structurally binds the entire *utenzi* together. This lexical choice, however, presents a notable philological curiosity. The vast majority of Swahili theorists, most notably Kaluta Amri Abedi, universally use the term *bahari* (ocean or sea) to describe this final constant rhyme. Given Shaaban Robert's undisputed status as a foundational master of the language, it would be presumptuous to dismiss this discrepancy as a mere typographical permutation of letters. Whether it reflects a deliberate conceptual nuance, a localized terminology, or an unresolved historical editorial anomaly, the text stands as written: for Shaaban Robert, the structural anchor of the *utenzi* remains the *habari*. Ultimately, what matters most for our present formalization is that this anchor operates purely as a structural parameter. Once again, the master poet defines the meter exclusively through an absolute syllabic grid and terminal homophony. There is no mention of exogenous accentual features, metrical feet, or stress patterns. The foundation of the verse remains strictly and syllabic.

5.2.3. The Epistemological Boundary: *Guni* as a Formal Defect

Perhaps the most devastating piece of evidence against those who would project loose, rhythmic models onto classical Swahili poetics is Shaaban Robert's definition of *Guni*. Within accentual Western paradigms, deviations from strict meter are often celebrated as

[31] Swahili original : "Utenzi ni shairi lenye habari katika mstari wa mwisho wa kila ubeti. Habari katika sanaa ya ushairi ni namna ya kina kinachofululiza kutokea katika kila mstari wa mwisho wa utenzi."

rhythmic variation or artistic license. In the classical Swahili matrix defended by Shaaban Robert, any failure to uphold the absolute syllabic count or the rhyme scheme is not a variation; it is a clinical defect.

He writes with uncompromising clarity: "Poems of *guni* are known by the defect of measure and the flaw in rhymes." (Robert, 1972, p. xiii)[32] The words chosen by the master—*dosari* (blemish, defect, error) and *ila* (flaw, defect)—are prescriptive and definitive. *Guni* is not an alternative artistic path; it is the breakdown of the formal contract. He further clarifies that even if a poem possesses syllables and rhymes, but violates the macroscopic structural rules (such as exceeding five lines per stanza without traditional precedent), it is instantly categorized as defective: "Poems which have arrangements of stanzas of more lines than five are *guni* just the same even though they have measure and rhymes." (Robert, 1972, p. xiii)[33]

Through *Sanaa ya Ushairi* [The Art of Poetry], Shaaban Robert confirms that the Swahili poetic universe is a closed, perfectly regulated system. He demands total submission to the endogenous laws of the *vina na mizani* system, sealing the argument that classical Swahili poetics yields to rhyme patterns, syllabic arithmetic, and formal rigor, not to imported theories of accentuation.

5.3. Ibrahim Noor Shariff: The Defense of Endogenous Sovereignty (1988)

Decades after Abedi, the scholar and master poet Ibrahim Noor Shariff solidified this formal consensus in his exhaustive 1988 work, *Tungo Zetu: Msingi wa Mashairi na Tungo Nyinginezo* (Our Compositions: The Foundation of *shairi* Poetry and Other Poetic Forms). Operating from a deeply patrimonial stance, Shariff fiercely defends the endogenous reality of Swahili metrics against foreign misinterpretations and academic projections.

Shariff explicitly claims sovereign ownership over the formal rules of Swahili versification. He opens his treatise with a prefatory *shairi* poem that asserts the endogenous authority of the Swahili formal

[32] Swahili original: "Mashairi ya guni hujulikana kwa dosari ya mizani na ila katika vina."

[33] Swahili original: "Mashairi yaliyo na mipango ya beti za mistari mingi kuliko mitano ni guni vile vile ingawa yana mizani na vina."

system, explicitly demanding that these rules be taught to those who remain ignorant of them (Shariff, 1988, p. ix):

> *Kitabu cha Tungo zetu, kina na uketo ndani*
> *Hakik'a hichi ni chetu, waungwana kisomeni*
> *Kimekusanya ya utu, na mambo ya kizamani*
> *Wasoyuwa wambiyeni, pop'ot'e kukit'apiya.*
>
> The book of our compositions, holds rhyme and depth within[34]
> Certainly this one is ours, oh noble ones read it
> It has gathered human qualities, and matters of antiquity
> Those who do not know, tell them, and distribute it everywhere.

By stating "Certainly this one is ours" (*Hakik'a hichi ni chetu*), Shariff rejects any exogenous attempt to redefine the metrics. The poetry and its formal laws belong to the Swahili tradition, and its mechanics must be understood from within.

He is acutely aware of the damage done by exogenous scholars who attempt to analyze Swahili poetry without understanding its true formal laws. He does not view these external analyses as mere academic isolated missteps, but as systemic, epistemological failures. He dedicates an entire section of his fifth chapter to this exact issue, naming it: *Upungufu katika utafiti na maelezo ya wageni* (Insufficiency in the research and explanations of foreigners) (Shariff, 1988, p. 161).

He specifies that this "insufficiency" is a profound technical failure. It stems directly from a fundamental inability to comprehend the strict syllabic and rhyming laws *(vina na mizani)* that construct the Swahili metrical grid. He titles the subsequent analytical subsection

---

[34] In his original Swahili text, Shariff employs a masterful semantic *double entendre* with the phrase *kina na uketo*. In Swahili poetics, the noun *kina* (Class 7, pl. *vina*) constitutes the fundamental metric lock of the rhyme, functioning as a structural and architectural boundary within the syllabic grid (see Chapter 4). However, in the standard lexicon, *kina* also signifies physical or conceptual "depth." The term *uketo* similarly translates to "depth" or "profundity." Thus, the pairing *kina na uketo* resonates simultaneously as the technical "rhyme and depth" and the literal "profound depth." The brilliance of the term *kina* lies in its double etymology: it represents a perfect conceptual convergence between an Arabic lexical loan (pertaining to rhyme) and a profoundly Bantu linguistic reality (the spatial root for depth). This dual identity perfectly encapsulates the endogenous nature of the Swahili metrical grid: it may assimilate external vocabulary, but its architectural logic and conceptual spatiality remain strictly and exclusively Bantu.

to pinpoint the exact cause of this foreign misinterpretation: *Upungufu unaotokana na kutojua arudhi* (Insufficiency resulting from not knowing metrics) (Shariff, 1988, pp. 161-182)

Building on Shariff's critique, the imposition of foreign rhythmic models—such as the search for accentual feet or temporal beats—is a definitive symptom of being blind to the real Swahili formal grid. Exogenous scholars project their own metric habits onto the texts precisely because they do not know how to read the endogenous syllabic laws that constitute Swahili metrics *(arudhi)*.

However, Shariff pushes this critique even further, elevating it from a technical dispute to a historical and ideological expose. He argues that this "insufficiency" in metrical analysis is not always an innocent mistake; it is entangled with a broader ideological, colonial, and religious prejudice. He illustrates how Western scholarship has historically sought to diminish, alter, or attack the Islamic Swahili heritage. He introduces this severe accusation under the heading: *Mfano wenye lengo la kupindua dini kwa kugeuza tarehe* (An example where the goal is to overthrow religion by changing dates) (Shariff, 1988, pp. 161-182).

To visually and viscerally demonstrate the depth of this exogenous colonial hostility toward the Islamic Swahili culture—which inevitably bleeds into how its literature and formal poetics are treated—Shariff analyzes a caricature from a British satirical journal (*Punch*) from 1889. He describes the explicit bias of the imperial gaze:

> Look at how the Western Christian states were drawn with beautiful shapes and appearances, and how the Muslim was drawn like a black dog with devil's horns, ready to be killed by cross-shaped swords. (Shariff, 1988, p. 36)[35]

By detailing this ideological distortion, Shariff proves that this specific exogenous lens is compromised. Extrapolating from his observation, this study identifies a highly opportunistic epistemological mechanism at play. Within this particular diffusionist framework, Islam and Arabness function as a malleable adjusting variable, molded at will to serve shifting ideological goals. When the objective is to undermine the faith itself—as in Shariff's historical

[35] Swahili original : "Angalia namna gani dola za Kizungu za Kikristo zilivyochorwa kwa maumbo na sura nzuri, na vipi Mwislamu ulivyochorwa kama mbwa mweusi mwenye pembe za shetani, tayari kuuawa na panga za msalaba."

example—the Islamic heritage is openly attacked. However, when the goal is to enforce —whether through conscious ideological design or the uncritical inheritance of a dominant colonial paradigm—what was defined in Section 1.4 as the "Exo-African Postulate," the artifacts of Islamic civilization are cynically weaponized to strip Swahili culture of its African identity.

This instrumentalization of the Swahili heritage is poignantly captured by the Lamu poet Ustadh Mau, who metaphorically describes his culture as a plaything subjected to the whims of outsiders (Roy & ibn Kadir, 2013):

Mimi ni kama mpwira	I [Kiswahili] am like a ball
Hutezewa uwandani	Played with on the field
Hupigwa teke ndiyani	Kicked along the way
Na kula anaepita	By every passerby

By framing any sophisticated structural element as an exclusively foreign, Oriental intrusion, this specific school of exogenous scholars methodically denies the endogenous Bantu genius behind Swahili creations. Consequently, the imposition of foreign rhythmic meters onto Swahili poetry operates as a direct metric avatar of the Great Zimbabwe Syndrome. Just as colonial historians attributed the stone architecture of Great Zimbabwe to mythical foreign builders, this specific exogenous gaze refuses to acknowledge the sovereign, endogenous formal law governing the Swahili syllabic grid, merely kicking the "ball" of Swahili metrics to fit their own theoretical playing field.

Having stripped away the noise of foreign misinterpretation and ideological prejudice, Shariff returns to the absolute reality of the Swahili poetics. A common critique leveled by exogenous theorists is that a strict syllabic and rhyming grid is inherently too rigid, suggesting that an underlying "rhythmic flexibility" or accentual beat must exist to allow for true poetic expression. Shariff completely dismantles this exogenous assumption. He argues that the absolute constraints of the syllable count (*mizani*) and the rhyme (*vina*) do not stifle the poet; rather, they provide the exact formal framework within which true artistic liberty is achieved.

> [...] there is no doubt that a wise composer will see that there is a great deal of freedom in composing with rhymes *[vina]* and syllabic measures *[mizani].* (Shariff, 1988, p. 48)[36]

Shariff's conception leaves no room for the ambiguity of accentual feet or temporal beats. The freedom he describes is not the freedom to break the meter, to insert rhythmic pauses, or to rely on a compensatory pulse. It is the absolute creative freedom to manipulate vocabulary within the strict, uncompromising boundaries of the exact syllable count and the rhymes. The formal syllabic grid is entirely sovereign, and the exogenous search for an accentual beat is nothing more than an academic construct born from "not knowing metrics" (*kutojua arudhi*).

5.3.1. The Lexical Architecture: *Vina*, *Mizani*, and *Kiwango* as an Indivisible Concept

Before outlining the specific spatial requirements of the stanzas, Shariff explicitly defines the technical lexicon of Swahili metrics (Shariff, 1988, pp. 42-44). What emerges from his definitions is an epistemological lock: the Swahili formal system cannot be conceptualized by divorcing the syllable from the rhyme or the caesura. They are the interlocking coordinates of a single spatial/mental and arithmetic reality.

Shariff establishes the absolute primacy of the syllable count as a literal act of calculation (*hisabu*):

> Measure *[Mizani]* is the count of the various sounds in every word. For example, the word *'hisabu'* [counting] has three measures: hi, sa, and bu. (Shariff, 1988, p. 42)[37]

He immediately locks this syllabic enumeration to the concept of the rhyme (*kina*), defining the rhyme not simply as an acoustic phenomenon, but as a strict formal boundary that only occurs "after a specific count of syllables".

---

[36] Swahili original: "[...] hapana shaka kuwa mtungaji mwenye busara ataona kuwa upo uhuru mwingi katika kutunga kwa vina na mizani."

[37] Swahili original: "Mizani ni hisabu ya sauti mbalimbali katika kila neno. Kwa mfano, neno 'hisabu' lina mizani tatu: hi, sa na bu."

> Rhymes *[vina]* are corresponding sounds contained in a stanza or a section of a composition; sounds which usually appear after a specific count of measures *[mizani]* and or at the end of every line.) (Shariff, 1988, p. 43)[38]

This definition is devastating for the accentual school. By defining the rhyme (*kina*) as an event that must occur *baada ya hisabu maalumu ya mizani* (after a specific arithmetic count of syllables), Shariff formally defines the Swahili metrical grid as a purely arithmetical and syllabic construct. According to his theory, the rhyme is the physical and mental marker that proves the syllable count was correct.

Even more, Shariff invalidates the exogenous concept of musical rests or free rhythmic pauses by formalizing the unrhymed caesura (*kiwango*). If a poet wishes to insert a pause without a rhyme, this pause is still ruthlessly governed by the exact syllable count:

> A *kiwango* [caesura/limit] is a stop after a specific syllable count *[mizani]* in every line of a stanza which does not have matching sounds or rhyme *[kina].* (Shariff, 1988, p. 43)[39]

The methodological imposition of the accentual school is clearly illustrated by how modern exogenous theorists attempt to reappropriate this specific endogenous definition. In a recent effort to sustain the accentual framework for archaic poetry, Emiliano Minerba quotes Shariff's definition of *kiwango* (the unrhymed caesura) accurately, and he grants its syllabic nature: "Either rhymed or unrhymed, however, the Swahili caesura is marked linguistically by the word boundary (besides being marked metrically by syllable count): a word cannot be placed across a caesura" (Minerba, 2025, p. 157). He also observes that the caesura "always appears in the same position (that is, after a certain number of syllables, which remains the same in all the lines)" (Minerba, 2025, p. 156). As a description, this agrees with Shariff.

The reinterpretation begins when Minerba turns from description to origins. To explain how the classical meters retained their syllabic

---

[38] Swahili original: "Vina ni sauti zilizolingana zilizomo katika ubeti au kifungu cha utungo; sauti ambazo aghlabu huonekana baada ya hisabu maalumu ya mizani na au mwisho wa kila mshororo."

[39] Swahili original: "Kiwango ni kituo baada ya mizani maalumu katika kila mshororo wa ubeti ambacho hakina sauti za milingano au kina."

constraints, he reverses the relation between stress and boundary: "the position of metric stresses determines the position of word-boundaries within the line" (Minerba, 2025, p. 213). The *kiwango* then becomes the trace of a vanished foot. In the classical age, he writes, "what remained of the earlier feet were sequences of syllables delimited by one or more caesurae, corresponding to word-boundaries," and "a systematic disposition of word boundaries corresponds to a systematic disposition of stresses, which reproduce the earlier archaic accentual metre patterns" (Minerba, 2025, p. 298). In the *wimbo*, the caesura receives a rhythmic purpose: it was inserted "to prevent the new classical metre from being articulated in trisyllabic, non-trochaic feet" (Minerba, 2025, p. 221).

This is the Procrustean bed in action. Whatever the merits of this account for the archaic songs (*utumbuizo*), which lie outside the scope of this book, the classical *kiwango* defined by Shariff needs no foot in order to exist. Its only condition is *mizani maalumu* (a specific syllable count) (Shariff, 1988, p. 43). The stress that precedes it is not its cause but its consequence: since the *kiwango* coincides with a word boundary, the penultimate stress of the word that closes the segment necessarily falls on the segment's second-to-last syllable. Minerba reads as a beat required by the meter what is only the reflex of a boundary placed by the count. If a syllable is missing from the prescribed count, the line is simply *guni* (defective); the *kiwango* compensates for nothing and serves only to delimit the units of the metrical grid.

Minerba's analysis overlooks the true nature of metrical flexibility in Swahili poetics. The freedom the poets describe belongs to the composer (*mtunzi*) during the architectural design of the poem: a poet is entirely free to construct a metrical grid where the *kiwango* falls after the fourth syllable, or alternatively, after the sixth or eighth syllable (Shariff, 1988, p. 45). However, once this formal parameter is selected for a specific poem, the *kiwango* becomes an absolute, unyielding metrical and syllabic wall. It is not a floating pause that a performer can shift or deploy to plug syllabic holes in a defective line; it is a rigid, invariable boundary dictated by the foundational laws of *vina na mizani*. Nothing in the data requires that this syllabic boundary be read as the vestige of an accentual foot; that reading is required only by the "trochaic innovation" that Minerba builds on the *Grenzschema* inherited from Vierke (2011) (see Sections 2.3 and 2.4). The formal integrity of the classical poem relies strictly on the arithmetic counting of syllables (*mizani*) and the spatial locking of

rhymes (*vina*) and caesuras (*viwango*), rendering any reliance on temporal pulses or exogenous musical accents completely void.

To permanently lock this arithmetic framework, Shariff defines the concept of *bahari* (the overarching meter) and *mkondo* (the specific sub-meter variations) (Shariff, 1988, p. 44). An accentualist might assume that a metrical variation implies a shift in "rhythm" or "beats." Shariff explicitly refutes this, defining variations exclusively as arithmetic shifts in the syllable count:

> The *bahari* [meter] is a single group of compositions that can have different sub-meters inside it [...] *Mikondo* are different kinds of compositions within one meter. The difference of *mikondo* is usually the difference in the count of measures *[mizani]*. (Shariff, 1988, p. 44)[40]

There is no space in this exhaustive formal taxonomy for a temporal "beat" or an accentual "foot." The generative engine is entirely accounted for by the arithmetic relationship between the syllable count *(mizani)*, the rhyme *(kina)*, and the measured pause *(kiwango)* (Shariff, 1988, pp. 42-44).

5.3.2. Shariff's Definition of the Classical Meters: *Utenzi*, *Wimbo*, and *Shairi*

Armed with these purely syllabic and arithmetic tools, Shariff proceeds to formally define the core meters of the classical Swahili tradition (Shariff, 1988, pp. 44-51). In every single definition, the exogenous search for an accentual foot is rendered obsolete by the sheer precision of the syllabic terminology. Our primary objective here is to give voice to Shariff to demonstrate exactly how he defines the meters—foundational definitions that the reader can directly compare with the structural models provided in the Quick Reference Guide and expanded upon in Chapter 8.

Following Shariff's own theoretical progression—beginning with the *wimbo*, moving to the highly complex *shairi*, and concluding with the epic *utenzi*—he strips away any performative melody. He reveals an uncompromising architectural blueprint based exclusively on exact line counts and internal syllabic divisions.

---

[40] Swahili original: "Bahari ni kumbo moja la tungo ambalo huweza kuwa na mikondo tafauti ndani yake [...] Mikondo ni aina tafauti ya tungo katika bahari moja. Tafauti ya mikondo huwa ni tafauti ya hisabu za mizani."

Starting with the *wimbo* meter, he explicitly categorizes its fundamental structure and its variations (*mikondo*) not by rhythmic beats or temporal musical phrasing, but strictly by their absolute line count, syllable totals, and spatial rhyme placement. He establishes the *wimbo* by its exact three-line architecture, before defining its internal variations as arithmetically bound units. He specifically highlights the standard "eight by eight" (*nane kwa nane*) syllabic measure as the canonical anchor of this rigid formal grid:

> A *wimbo* is a stanza of three lines which can have different variations. [...] The element that distinguishes the *wimbo* meter from other meters is the exact count of three lines in each of its stanzas. [...] But the variation of eight by eight syllables, that is, having a rhyme on the eighth and sixteenth syllable in each line, is the most well-known. (Shariff, 1988, p. 45)[41]

He then establishes the canonical standard of the *shairi* by its spatial alignment, insisting once again on the strict arithmetic rule of the *mizani* rather than any exogenous temporal metric footprint. Here, the generative core relies entirely on the precise syllabic count across the hemistichs, locking the stanza into an unyielding metrical grid and leaving no theoretical void for an alien stress pattern:

> A *shairi* is a stanza of a composition of four lines or more with internal rhymes [...]. Likewise, eight by eight syllables per line is the most favored. (Shariff, 1988, p. 49)[42]

Finally, regarding the epic *utenzi*, Shariff describes the spatial layout entirely through the mandatory placement of rhymes (*vina*)

---

[41] Swahili original : "Wimbo ni ubeti wa mishororo mitatu ambao unaweza kuwa na mikondo tafauti. [...] Kitu kinachopambanua bahari au kumbo la nyimbo kutokana na bahari nyingine ni idadi ya mishororo mitatu katika kila ubeti wake. [...] Lakini mkondo wa mizani nane kwa nane, yaani yenye kina katika mizani ya nane na ya kumi na sita katika kila mshororo, ndiyo yenye kujulikana zaidi."

[42] Swahili original : "Shairi ni ubeti wa utungo wa mishororo mine au zaidi ya mine wenye vina kati [...]. Vilevile mizani nane kwa nane kila mshororo ndiyo inayopendwa sana."

and the absolute line structure, which the composer actively sets as the foundational formal and syllabic grid:

> An *utenzi* or *utendi* is a single stanza of a composition consisting of two lines *(mishororo)* containing three rhymes: in the middle and at the end of the first line, and in the middle of the second line. The sound of the final rhyme can be any other sound the composer desires. Many composers since ancient times preferred that the final sound of every stanza be identical, thus also forming a rhyme. (Shariff, 1988, p. 51)[43]

At this juncture, a clarification is necessary for readers accustomed exclusively to romanized *utenzi* editions. While modern transliterations typically stack the stanza into four distinct graphic lines, Shariff's definition strictly relies on the *mshororo*—the formal metrical line intrinsic to the endogenous Swahili matrix and the Arabic-script *Ajami* manuscripts (as defined in Chapter 4). In this authentic framework, the stanza consists of precisely two lines (*mishororo*), each bisected into two segments. Consequently, what the romanized layout presents as a vertical column of four short lines is structurally engineered as two tightly bound horizontal lines. The first three segments share the same rhyme, while the fourth segment carries the overarching terminal rhyme that unifies the entire composition.[44] Shariff provides a highly valuable structural nuance regarding this final segment: the omnistrophic terminal rhyme is technically optional. He explicitly notes that the sound of this final rhyme remains entirely at the composer's sovereign discretion, even

---

[43] Swahili original : "Utenzi au utendi ni ubeti mmoja wa utungo wa mishororo miwili uliokuwa na vina vitatu: kati na mwisho wa mshororo wa kwanza na kati ya mshororo wa pili. Sauti ya kina cha mwisho huwa nyingine yoyote apendayo mtungaji. Watungaji wengi tokea zama za zama, walipendelea sauti ya mwisho ya kila ubeti iwe moja, iwe vina pia."

[44] It should be noted that despite this structural reality, the spatial matrix formalized in this study (see QRG.2 for the *utenzi* meter) utilizes the modern romanized presentation. This decision reflects the material reality of the available corpus: the stacked romanized layout overwhelmingly dominates both contemporary Swahili publications and modern printed editions of 19th-century classical texts. Furthermore, original Arabic-script *Ajami* manuscripts are often difficult to access. As a direct consequence of colonial-era looting, many of these vital primary sources remain scattered across Western archives in Germany and the United Kingdom. For these historical and editorial reasons, the formal models developed in this study focus exclusively on the accessible romanized corpus.

though maintaining a single unifying rhyme across the entire poem has been the overwhelmingly dominant practice since ancient times.

Across all these classical genres, the Swahili poet (*mtunzi*) engineers the physical mass of the line using exact arithmetical and syllabic measures, firmly securing the verses within a strict formal and syllabic grid. By explicitly defining the metrics through syllables and exact calculation, the endogenous theory completely disregards any phantom rhythmic pulse, proving that the accentual framework is a coercive exogenous projection rather than an intrinsic reality of the Swahili poetic matrix.

To conclude this demonstration and confirm that the *kiwango* is always a fixed syllabic boundary, we must examine Shariff's analysis of *Al-Inkishafi* (Shariff, 1988, pp. 54-55). An apocalyptical[45] poem composed by Sayyid Abdallah bin Ali bin Nassir, *Al-Inkishafi* is universally regarded as a supreme masterpiece of classical Swahili poetry. Shariff classifies its meter under the *Dura Mandhuma / Inkishafi* category and provides a clinical, arithmetic breakdown of its internal syllabic grid. He explains exactly how the unrhymed caesura (*kiwango*) functions within its specific 11-syllable line:

> This meter has stanzas of four lines having measures *[mizani]* of six by five and its rhymes are at the end of the three lines of each stanza, therefore the sixth measure is just a *kiwango*. (Shariff, 1988, p. 55)[46]

This single sentence dismantles the accentualist framework. Shariff does not describe the *kiwango* as a pause to catch a breath, nor as a temporal beat compensating for a rhythmic syncopation. He defines the *kiwango* exclusively by its spatial and numerical coordinate:

---

[45] We use the term apocalyptical because it mirrors the exact etymological essence of the title *Inkishafi*. The Swahili term derives from the Arabic root *kashafa* (كشف), meaning "to uncover" or "to reveal" (which notably yields the Swahili noun *kashifa*, referring to a scandal or the exposing of a hidden disgrace). The Arabic verbal noun *inkishāf* (اِنْكِشاف) literally translates to "an unveiling" or "a revelation." This conceptual matrix corresponds flawlessly to the Ancient Greek verb ἀποκαλύπτω (*apokaluptō*), formed from *apo-* (away) and *kaluptō* (to cover), meaning precisely "to uncover" or "to reveal." Across these three linguistic traditions, this process describes not merely a catastrophic end, but the profound unveiling of an ultimate truth wherein worldly illusions collapse.

[46] Swahili original : "Bahari hii ina beti za mishororo minne yenye mizani sita kwa tano na vina vyake viko mwisho wa mishororo mitatu ya kila ubeti, kwa hivyo mizani ya sita ni kiwango tu."

mizani ya sita (the sixth syllable). The line is strictly divided into an asymmetrical 6+5 syllabic grid. Because the rhyme (*vina*) only occurs at the end of the line, the boundary explicitly anchoring the end of the first six syllables is an unrhymed caesura—a *kiwango*.[47]

To observe this strict formal grid in practice, we rely on the verses exactly as cited by Shariff (1988, p. 55). To accurately render the epic tone of the text for readers unfamiliar with Swahili, we have paired Shariff's Swahili transcription with the classic English translation from William Hichens' 1939 edition, *Al-Inkishafi: The Soul's Awakening:*

Dunia ni jifa siikurubu
Haipendi mt'u ila kilabu[48]
I-hali gani ewe labibu
Kuwania mbwa hutukizwaye!

The world is a carrion corpse, do not approach it;
No man loves it, except dogs;
What state of mind is this, O man of understanding,
To wrangle for it with dogs, are you not disgusted?

---

[47] As explored in the terminological definition of the *kiwango* in Chapter 4, it is precisely from this strict structural reality established by Shariff that we formalized the theoretical hypothesis of this unrhymed boundary operating fundamentally as a "rhyme zero" [*rime zéro*] (Roy, 2013b, pp. 305, 313, 1110).

[48] While we retain Hichens' (1939) canonical translation in the main text, his reading of the second line ("No man loves it, except dogs") relies on a syntactic parsing that creates a somewhat disjointed semantic structure. Hichens interprets *haipendi* with a Class 1 negative subject prefix (*ha-*) and a Class 9 object infix (*-i-*), making *mt'u* (man) the subject. This forces "dogs" (*kilabu*) to act as a contrasting subject (i.e., men do not love the world, but dogs do).

However, classical Swahili morphology allows for a more elegant reading. Parsing *haipendi* with a Class 9 negative subject prefix (*ha-i-*), referring directly back to *dunia* (the world), yields: "It [the world] loves no man, except dogs." This creates a perfectly balanced structure where both "man" and "dogs" are objects of the world's disposition. Given the first line's metaphor of the world as a carrion corpse (*jifa*), it is poetically far more cohesive to state that the world rejects humanity and reserves its affection strictly for scavengers.

A strict morphologist might object that if the animate noun *mt'u* is the object, the verb should technically feature the Class 1 object infix *-m-* (yielding *haimpendi*). Yet, in classical Swahili poetry, such grammatical markers are routinely elided to satisfy the uncompromising syllabic grid. Inserting the object infix would generate seven syllables (*ha-i-m-pen-di m-t'u*), shattering the absolute six-syllable boundary of the *kiwango*. By eliding it, the composer perfectly locks the first segment into the mandatory six syllables (*ha-i-pen-di m-t'u*), seamlessly followed by the five syllables of the second

Kamwe ina ila iliyo mbovu
Ilikithiriye ungi welevu
Ikalifu mno kutaka mavu
Kupa wat'u ngeya ikithiriye

It has nothing but vile defects;
It abounds in profound deceit;
It is excessively fierce in seeking ruin;
In giving men strife, it has multiplied.

Sayyid Abdallah bin Ali bin Nassir, *Al-Inkishafi*, stanzas 19–20[49]

The *kiwango* is an immovable metrical wall built from syllabic bricks, fixed at the exact boundary of the sixth syllable (*mizani ya sita*) (Shariff, 1988, p. 55). By isolating and mapping the exact syllabic coordinates of this stanza, the arithmetic and syllabic supremacy of the grid becomes undeniable:

Ka-mwe i-na i-la (6) [*kiwango*] i-li-yo mbo-**vu** (5) [rhyme vu]
I-li-ki-thi-ri-ye (6) [*kiwango*] u-ngi we-le-**vu** (5) [rhyme vu]
I-ka-li-fu m-no (6) [*kiwango*] ku-ta-ka ma-**vu** (5) [rhyme vu]
Ku-pa wa-t'u nge-ya (6) [*kiwango*] i-ki-thi-ri-ye (5)

This exact count proves that the *kiwango* is not a freely placed temporal pause used to compensate for a missing syllable within an accentual foot—contrary to the framework proposed by Minerba—but rather a rigidly fixed, structural barrier.

In Shariff's theoretical framework, the Swahili poetic intellect operates through the absolute counting of syllables (*mizani*)[50], which subsequently defines the precise spatial boundaries of the line (*vina* and *kiwango*) (Shariff, 1988, pp. 43, 55). This endogenous syllabic system is entirely self-sufficient, rooted in an architectural vision of the text rather than a temporal one dictated by a regular beat or rhythmic pulse. To project a Western musical accent onto this flawless

segment (*i-la ki-la-bu*). This demonstrates how the rigorous formal and numerical constraints of the *mizani* supersede standard prose morphology.

[49] Swahili text from Shariff, 1988, p. 55; English translation from Hichens, 1939.

[50] A fundamental distinction must be drawn between this formal metrical analysis and the practical act of creation. While we analyze the structural grid here in strictly arithmetic and syllabic terms, the traditional Swahili poet relies on a deeply internalized cognitive mastery of these boundaries. They neither mechanically count syllables on their fingers, nor do they scholastically fill out a blank metrical matrix on paper. For a detailed exploration of this cognitive process in practice, see Section 3.1.

arithmetic and syllabic architecture is not merely an analytical error; it is to ignore the explicit instructions of this master poet and scholar. By forcing the classical Swahili verse into a Procrustean bed of exogenous metrics, the accentualist approach actively dismantles the rigorous formal matrix that the *mtunzi* painstakingly engineered.

5.4. Ahmed Sheikh Nabhany and the Exhaustive Taxonomy of Classical Forms (2006)

This strict formal logic finds its most exhaustive and culturally authoritative articulation in the work of the late Sheikh Ahmed Sheikh Nabhany, widely regarded as one of the supreme guardians of classical Swahili poetics. Nabhany's theoretical framework is crystallized in his detailed *Dibaji* (Foreword) to K. W. Wamitila's *Kamusi ya Ushairi* (Dictionary of Poetry), published in 2006.

In this foundational text, Nabhany systematically outlines the formal requirements of classical Swahili verse, starkly contrasting the highly regulated ancient forms against modern, unregulated prose poetry. His definition of poetry rests on absolute formal laws that the composer must obey:

Many of the ancient compositions, which this author has studied, were composed following principles and laws—not by writing prose *(nathari)* as happens in modern poetry.

That poetry adhered to rules such as the meter *(bahari)*, the rhyme *(kina)*, the line *(mshororo)*, the submeter *(mkondo)*, the caesura *(kituo)*, the conclusion *(kikomo)*, the [syllabic] measure *(mizani)*, and the limit *(kiwango)*. (Nabhany, 2006, p. [3])[51]

When defining the rules of composition, Nabhany does not actively argue against the concept of an accentual "foot"—because such an exogenous notion simply does not exist within the authentic cognitive framework of the *mtunzi*. Instead, he presents the endogenous generative engine of the poetry purely on its own structural terms. Operating strictly from within the living tradition as an

[51] Swahili original : "Tungo nyingi za kale, ambazo amezisoma mwandishi huyu, zilitungwa kwa kufuata kanuni na sharia- sio kwa kuandika tutumbi (nathari) kama inavyotokea katika ushairi wa siku hizi. Ushairi huo ulizingatia kaida kama: bahari, kina, mshororo, mkondo, kituo, kikomo, mizani na kiwango."

uncompromising purist, he establishes an ultra-strict, hermetically sealed taxonomy. He states that the entire classical corpus is categorized into exactly thirteen distinct formal categories (*bahari*):

> *Bahari* za tungo za Kiswahili ni kumi na tatu, nazo ni (The poetic meters of Swahili compositions are thirteen, which are):
> (a) *Shairi*
> (b) *Wimbo*
> (c) *Utendi*
> (d) Dura Mandhuma/Inkishafi
> (e) Ukawafi
> (f) Tiyani Fatiha
> (g) Hamziya
> (h) Utumbuizo
> (i) Wawe/Vave
> (j) Kimai
> (k) Zivindo
> (l) Sama/*Mahadhi*/sauti.
> (Nabhany, 2006, pp. [3]-[4])

A rigorous, data-driven reading of this foundational primary source reveals a fascinating textual anomaly. Out of profound reverence for Nabhany's unmatched authority, the academic community has historically taken his number "thirteen" for granted, likely lulled into complacency by the alphabetical enumeration. However, a strict count of the data reveals that the list concludes at the letter (l), containing only *twelve* items. This striking omission invites a compelling hypothesis regarding the unlisted thirteenth element: *guni*.

In classical Swahili poetics, *guni* refers exclusively to defective poetry—verses that fail to meet the uncompromising arithmetic and syllabic boundaries—rather than a distinct, legitimate poetic genre. For a structural purist like Nabhany, who famously exhibited zero tolerance for metric deviation or free verse, *guni* conceptualizes the absolute boundary of what is acceptable. It is highly probable that he conceptually numbered *guni* as the thirteenth category of composition, but his deep-seated purism caused him to unconsciously—or perhaps deliberately—omit it from the physical list of valid, noble *bahari*.

5.4.1. The Three Primary Models: *Shairi*, *Wimbo*, and *Utendi*

Of these thirteen professed meters—though, as we have established, only twelve actually materialize on the page—Nabhany isolates three as the most prominent, widely utilized, and foundational to the tradition. The remaining meters, he notes, demand extraordinary expertise:

> Among the meters that are famous, widely used, and not as difficult as the others are: 1. *Shairi*, 2. *Wimbo*, 3. *Utenzi/Utendi*. The remaining genres are difficult and require a very highly skilled poet to compose. (Nabhany, 2006, p. [4])[52]

Nabhany then proceeds to define the exact formal rules governing these three core forms. His clinical definitions are anchored in a strictly arithmetic and syllabic foundation, entirely corroborating the endogenous framework championed by Shariff. He defines the poetry through absolute syllable counts (*mizani*) and structural boundaries like rhymes *(vina)*, leaving no room for exogenous rhythmic beats or accents.

1. Shairi

Nabhany defines the *shairi* meter by its strict four-line stanza and its rigid middle and terminal cross-rhyming scheme.

> The *Shairi* [meter] consists of four lines. The middle rhymes align, and the terminal rhymes likewise, but in the final line, the terminal rhyme come to the middle, and the middle rhyme go to the end. (Nabhany, 2006, p. [4])[53]

[52] Swahili original : "Katika bahari ambazo ni maarufu na zatumiwa na si ngumu kuliko nyingine ni: 1. Shairi 2. Wimbo 3. Utenzi/Utendi. Bahari zilizosalia ni ngumu na yataka mshairi hodari sana kuweza kutunga."

[53] Swahili original : "Shairi ni mishororo minne. Vina vya kati hulingana vya mwisho kadhalika lakini mshororo wa mwisho kina cha mwisho huweza kuja kati na cha kati kikaja mwisho."

2. Wimbo

He similarly codifies the *wimbo* meter strictly by its stanza length and its rhyming obligations, cementing its status as a formal blueprint rather than a loose song.

> The *Wimbo* [meter] consists of three lines, the middle and terminal rhymes must align. (Nabhany, 2006, p. [4])[54]

3. Utendi (Utenzi)

For the *utendi* meter, Nabhany provides a masterful breakdown. He highlights its narrative function before detailing its precise formal dimensions, explicitly identifying two distinct spatial instantiations (horizontal and vertical) that share the exact same underlying syllabic logic.

> As for the third meter, it is the *utendi*, the goal of which is usually composed to explain the history of a hero or champion, the history of a place, or counsel, or other topics that are developed at length. *Tendi* usually vary in length. [...] *Tendi/tenzi* can have two different [formal] arrangements: (a) two lines which have a single rhyme across three segments, but the final one is different; (b) four lines—each of eight syllables *[mizani minane]*—with the rhymes of the first [three] lines matching, and the final one *(bahari)* being different. (Nabhany, 2006, p. [5])[55]

Unlike Shariff, who focuses his definition on the horizontal layout (two lines of sixteen syllables divided into two *vipande*), Nabhany explicitly acknowledges both visual arrangements of the same arithmetic reality. Arrangement (a) refers to the horizontal spatial instantiation, while arrangement (b) refers to the vertical one. A horizontal 16-syllable line (*mshororo*) is divided into two 8-syllable segments. Consequently, a two-line stanza yields exactly four 8-

---

[54] Swahili original : "Wimbo ni mishororo mitatu, vina vya kati na vya mwisho vilingane."

[55] Swahili original : "Ama bahari ya tatu ni utendi ambao lengo huwa unatungia kueleza mapisi (historia) ya mtu shujaa au jagina, mapisi ya pahala au wasia au maudhui mengine yanayoendelezwa kwa urefu. Tendi huwa na urefu tofauti. [...] Tendi/tenzi huweza kuwa na mipangilio miwili tofauti: (a) mishororo miwili ambayo huwa na kina kimoja katika vipande vitatu lakini kile cha mwisho kikawa tofauti, (b) mishororo minne- minne-ya mizani minane- vina vya mistari ya kwanza vikilingana na kile cha mwisho (bahari) kikiwa tofauti."

syllable segments, perfectly mirroring the formal parameters of the four-line vertical arrangement. Both instantiations will be fully modeled and formalized in Chapter 8, although the graphical matrices developed in this study adopt the vertical layout[56].

This dual layout reflects the complex material history of Swahili literature. The horizontal arrangement is tied to the *Ajami* script (Swahili written in modified Arabic characters), whereas the vertical arrangement aligns with the romanized script. However, one must avoid the simplistic chronological dichotomy equating *Ajami* with "ancient" and romanized with "modern." The *Ajami* script remains a living practice among contemporary poets. Conversely, the vast majority of the ancient classical corpus studied today is accessed through romanized transliterations. This reliance on transliterated text is a direct legacy of the colonial era, during which countless original *Ajami* manuscripts were looted and displaced primarily to European archives that are not yet fully digitized and accessible. Consequently, the romanized vertical layout—widely disseminated through printed editions—has become the standard vehicle for transmitting this heritage.

Regardless of the spatial layout chosen—horizontal *Ajami* or vertical romanized—the fundamental principles of the meter remain identical. The first of these foundational principles is the chain of rhymes: the internal rhymes linking the first three segments of the stanza, culminating in the final recurring rhyme (*bahari*). These precise rhyme boundaries are in turn strictly locked into place by the second core principle: the *mizani* (syllable counting).

This definition is strong testimony against the accentual school. By explicitly stating that the core generating mechanism relies on the specific parameter of *mizani minane* (exactly eight syllables), Nabhany confirms the endogenous reality of the syllable count. The formal structure of the meter is dictated by this eight-syllable boundary. Because the utendi is built by fitting words into these exact arithmetic limits, any exogenous search for an underlying accentual foot is superfluous; whether such a foot is nonetheless present in the

[56] True to our data-driven methodology, this is a practical choice: the vast majority of readily accessible texts today are romanized transliterations, which conventionally utilize Nabhany's vertical arrangement (b). To guide readers through this highly technical analysis, the definitive matrices and their corresponding algebraic equations—formalized in Chapter 8 and tested for robustness against the empirical data in Chapter 9—have been conveniently compiled at the very beginning of this book in the Quick Reference Guide.

lines is a question for the empirical tests of Chapter 9. The architecture of the utendi is generated by a sovereign numerical and syllabic grid, not by a temporal beat.

5.5. Hashil Seif Hashil: Acknowledging the Boundaries and the Formal Rupture (2009)

To prove that the classical Swahili metrical system is an uncompromising formal and syllabic grid—and not a flexible rhythmic tendency—we must examine how modern masters continue to engage with it. The late poet Hashil Seif Hashil, who resided between Denmark and Tanzania, demonstrates that *vina na mizani* is a sovereign generative matrix rather than a closed, static historical artifact.

The discourse surrounding Swahili poetry often falls into a simplistic binary opposition between the "ancients" (strict traditionalists) and the "moderns" (advocates of total rupture and free verse). However, this dichotomy fails to account for a sophisticated third path: the *ex materia* creation. Hashil Seif operates within this integrative mode of composition, unsettling both the rigid traditionalists and the radical moderns by exposing the limits of defining oneself merely by opposition.

To fully grasp the epistemological weight of this *ex materia* approach, one must draw an absolute distinction between the deep principles of the Swahili poetic system and its surface parameters. The foundational principles are immutable: they are the structural necessity of rhyme chains (*vina*) and the strict rule of absolute syllabic measurement (*mizani*). The parameters, conversely, are the specific numerical and structural configurations assigned to these principles. An *ex materia* innovator like Hashil Seif does not abandon the deep principles; rather, he forges his own metrical innovations through derivations and deliberate transformations of the pre-existing parameters.

In the introduction to his 2009 anthology *Kurunzi* (*The Torch*), Hashil directly addresses the traditionalists, advocating for a neutral space where formal creation remains open. He explicitly acknowledges the foundational principles while vehemently rejecting the dogmatic paralysis of blind obedience to frozen rules.

> The masters of writing ancient and contemporary poetry are guided by following specific rules and procedures of writing. For example – a poem must be able to be sung like a song, and also it must have rhymes. [...] Personally, I do not oppose their procedures of composition, what I ask is that we do not try to place obstacles in writing poetry, like traffic laws: if you see a red light stop, a green light continue your journey. People should write in the way they see fit. [...] Let us not make poetry like a religion that cannot change. (Hashil, 2009, p. 10)[57]

Hashil's argument is formalist, and his traffic light metaphor must be understood precisely. He does not prohibit strict traditionalists from composing classical verses; indeed, he explicitly tolerates and respects their procedural choices. What he vehemently rejects, however, is being subjected to their normative injunctions. This nuance is essential when evaluating the role of a strict conservator like Nabhany. Nabhany's unwavering rigidity is immensely valuable when defending Swahili poetry against exogenous theoretical attacks—such as the imposition of Western accentual metrics—because it fiercely protects the endogenous heritage. Yet, within the internal sphere of poetic creation, this exact same conservatism transforms into a compositional prison when the parameters are enforced as a fixed, unalterable dogma. Hashil's warning against treating poetry as an inviolable religion is essentially a demand for the right to formal "apostasy." He argues against the robotic, unthinking submission to a system treated as absolute law. If classical metrics represent the road, Hashil claims the right to re-engineer its structural parameters. Just as a society might shift its traffic laws without destroying the concept of driving, a poet must be able to redesign the arithmetic and topological configurations of a poem without facing excommunication from the poetic tradition.

To understand true *ex materia* innovation, we must correct a common misconception regarding traditional parameters. For instance, creating a six-line stanza or utilizing asymmetrical

---

[57] Swahili original : "Mabingwa wa uandishi wa mashairi wa kale na wa kisasa, wanaongozwa kwa kuf[u]ata kanuni na taratibu maalumu za uandishi. Kwa mfano – shairi lazima liweze kuimbika kama nyimbo, na pia liwe na vina. [...] Binafsi sipingi taratibu zao za utungaji, ninaloomba ni kwamba tusijaribu kuweka vipingamizi vya uandishi wa mashairi, kama vile sharia za barabani : ukiona taa nyekundu simama, taa ya kijani endelea na safari yako. Watu inafaa waandike kwa namna wanavyo ona wao ni sawa. [...] Tusifanye ushairi kama dini kwamba hauwezi kubadilika."

hemistiches are not inherently modern inventions; the classical matrix already accommodates such variations. One only needs to look at the hyper-fractionated *wimbo* matrix documented by Ibrahim Noor Shariff, which features three lines deliberately cut into four asymmetrical hemistiches. The traditional matrix is inherently elastic. However, a true modern generative operation goes much further into structural novelty. A contemporary poet might architect an entirely unprecedented stanza: for example, opening with two 10-syllable lines (each perfectly divided into 5+5 hemistiches), followed by a solid, undivided 8-syllable line without any hemistich, and closing with another two 10-syllable lines (5+5). This specific architecture does not exist in the classical corpus, yet it remains completely governed by the deep principles of *vina na mizani*—an absolute syllable count and rhyming homophony defining the boundaries of the hemistiches and lines.

This radical variation of parameters has a profound epistemological consequence: it pushes the poetry firmly into a strictly graphic dimension, acting much like the visual spacing of modern verse. Classical Swahili poetry is traditionally tied to oral performance, structured around a vital contract between two distinct entities: the poet-architect (*mtunzi*), who provides the rigid, written arithmetic and syllabic grid, and the singer (*mghani*), who performs it. The singer depends on the predictability of the classical syllables to lay down their melody (*mahadhi*), which operates independently of natural speech rhythms. By engineering a highly irregular stanzaic architecture or an entirely new one, the modern poet deliberately severs this performative contract. Confronted with an unprecedented structural map, the singer can no longer anticipate the grid, rendering the poem impossible to sing in the traditional manner. The poem achieves a sovereign spatial freedom for the reader, existing as a purely visual and intellectual object on the written page, yet it remains rigorously measured poetry.

By pleading for this aesthetic openness, Hashil warns that rejecting formal parameter variations constitutes an obstacle to the normal

evolution of cultural forms and the emergence of the new poet. He anticipates the possibility of a complete formal rupture:

> It is not at all impossible that in the coming days a new writer of poetry might appear. Perhaps we will also abandon our ancient and modern styles, bury them, and follow new styles that will emerge. (Hashil, 2009, p. 10)[58]

To allow for this emergence, Hashil demands the fundamental right to experimental exploration:

> But this cannot happen if we do not allow people to write poetry in diverse ways. (Hashil, 2009, p. 10)[59]

The verb he chooses, "to bury them" (*kuizika*), underscores the finality of the formal rupture, distinguishing it from his own transformative practice. Hashil confirms that the *vina na mizani* system is an exact, formal law that can be deliberately manipulated by the architect of the page. This deliberate manipulation of the formal grid—altering the parameters while fiercely guarding the principles—is precisely the path pioneered decades earlier by Mathias E. Mnyampala. Acting as both an ancient and a modern, Mnyampala anchored his work in the absolute arithmetic and syllabic reality of *vina na mizani* to compose both strictly classical verses and groundbreaking neo-classical forms.

5.6. Mathias E. Mnyampala: The Generative Expansion of the Formal Syllabic Grid (1963–1971)

While scholars like Ahmed Sheikh Nabhany viewed the classical Swahili metrical system as a closed, exhaustive taxonomy of a strictly delimited number of meters, Mathias E. Mnyampala proved that the *vina na mizani* grid is, in reality, a limitless generative matrix.

[58] Swahili original : "Si hasha kwa siku zijazo akatokea mwandishi mpya wa uandishi wa mashairi. Huenda pia tukaacha mitindo yetu ya kale na ya kisasa, kuizika na kufwata mitindo mipya itakayochomoza."

[59] Swahili original : "Lakini jambo hili haliwezi kutokea ikiwa hatuto waruhusu watu kuandika mashairi kwa njia mbali mbali."

Mnyampala, a master poet, historian, and magistrate, did not view the classical formal rules as a dead artifact or an inviolable religion. Instead, he treated the absolute syllable count and the rhyme as a sovereign foundation upon which new formal configurations could be built.

Mnyampala's approach to formal poetics perfectly mirrors his philosophical approach to the construction of the Tanzanian nation. In his biography of the Gogo chief, Mtemi Mazengo—an unpublished manuscript recovered from the Dodoma archives, the photographic facsimiles of which are preserved under the archival code mmd1 (Roy, 2013b, pp. 971–974)— he utilizes a metaphor of construction that applies equally to his vision of the poetic grid:

> A solid house was built upon a foundation made strong by skillful and ingenious artisans. Thus, it is exactly the same with a nation, which cannot be forged if there are no ethnic groups. The ethnic groups of the country are the foundation of the nation.[60]

Just as the nation is built upon the solid foundation (*msingi*) of the local ethnic groups, Mnyampala's poetic innovations are built directly upon the unyielding foundation of the classical *vina na mizani* rules. He does not erase the foundation; he utilizes it to erect new forms.

Mnyampala possessed a flawless mastery of the classical 16-syllable *Shairi* formal grid (divided strictly into 8+8 hemistiches). He did not rely on any exogenous accentual beat; he relied entirely on the exact syllable count and the spatial alignment of the rhymes. This perfection is evident in his poem *Ulimwengu ni Sarabi* (The World is a Mirage), published in his *Diwani* (1963):

> *Ulimwengu si pungufu, hubana vidhabidhabi,*
> *Huangusha watukufu, huyeyushwa kwa kalibi,*
> *Daima ni badilifu, kama kanga za mabibi,*
> *Ulimwengu ni sarabi, humeta na kumetua.*
>
> The world is not deficient, it squeezes the slanderers,
> It brings down the glorious, it is melted in a crucible,
> It is constantly changing, like the kanga cloths of ladies,
> The world is a mirage, it sparkles and fades.

---

[60] Swahili original : "Nyumba imara ilijengwa juu ya msingi uliofanywa imara na mafundi hodari na mahiri. Ndivyo vivyo hivyo taifa haliwezi kuundika ikiwa hakuna makabila. Makabila ya nchi ndio msingi wa taifa."

Tamati ukamilifu, anayeleta ni Rabi,
Mkidhi wetu Raufu, Muumba wetu Hababi,
Atakalo husadifu, sinibishie muhebi,
Ulimwengu ni sarabi, humeta na kumetua.

In the end, perfection, the one who brings it is the Lord,
Our Most Gentle Provider, our Beloved Creator,
Whatever He wills is realized, do not contradict me my friend,
The world is a mirage, it sparkles and fades.
(Mnyampala, 1963, pp. 115-116)

Mnyampala seamlessly combined his flawless classical mastery with a drive for systematic structural innovation. Throughout his body of work, he demonstrated the capacity to compose strictly classical verses while concurrently re-engineering the parameters of the metrical grid to forge unprecedented poetic forms. As detailed by his son Charles Mnyampala in the foreword to *Maisha ni kugharimia* (Mnyampala, 2013b, pp. xi-xxvi), the poet actively expanded the traditional architecture to create new configurations—most notably the *Mtiririko*, the *Msisitizo*, and the *Ngonjera*. The sequence of the formal examples presented below directly follows Charles Mnyampala's exposition of his father's creations, while the accompanying analysis unpacks exactly how the poet manipulated the deep arithmetic and topological grammar of the classical matrix.

5.6.1. The *Mtiririko* (The Flow)

It operates as a rigorous structural extension of the classical *shairi* matrix. Mnyampala preserves the absolute arithmetic of 16 syllables per line split into symmetric 8-syllable hemistiches. He also perfectly preserves the classical topological mapping of the rhyme chains (*vina*). The poem employs the standard cross-rhyming permutation (an "X" pattern) technically termed *kupindua* where the rhyme of the first hemistiches in lines one to three crosses over to the second hemistich of line four, and vice versa. Because this topology adheres strictly to the orthodox definition of the *shairi* genre recognized by purists like Nabhany, the structural innovation does not lie in the directional flow itself.

Instead, the generative breakthrough of the *Mtiririko* lies entirely in the unbroken scope and phonological permanence of the rhyme chains. In a standard *shairi*, the specific phonetic values of the rhymes are reset from one stanza to the next. In Mnyampala's *Mtiririko*—as

explicitly demonstrated in the poem *Tungeni mtiririko* (Mnyampala, 2013b, pp. xiii-xv)—the exact same phonetic values (bu / ko), are maintained continuously across the entire ten-stanza framework. By locking the sound values globally rather than locally, Mnyampala creates a steady, unbroken current of homophony that demands exceptional lexical mastery to prevent the linguistic flow from running dry.

5.6.2. The *Msisitizo* (The Emphasis)

Mnyampala pushed this parameter manipulation even further with the *Msisitizo*, as shown in the poem *Natunga msisitizo* (Mnyampala, 2013b, pp. xvi-xix). Here, the stanzaic architecture expands into a highly complex six-line hybrid structure that uses the classical *shairi* meter as its foundational base while radically altering its spatial and visual parameters.

The topological mapping of the rhymes relies on a specific classical variation known as *kupindua kina kimoja* (the inversion of a single rhyme). In this model, the cross-rhyming does not form a complete "X". Instead, only one rhyme chain—the one originating on the right side (the second hemistiches)—shifts. This right-hand rhyme cascades vertically down the stanza before suddenly crossing over to the left side (the first hemistich) exclusively on the sixth and final line, which serves as a fixed refrain.

However, Mnyampala introduces two unprecedented generative innovations to this matrix.

The first major rupture occurs on the third line of every stanza. Mnyampala deliberately breaks the classical 16-syllable symmetry by completely amputating the first hemistich. The third line is thus reduced to a single, isolated octosyllabic segment. Because it strictly carries the rhyme of the right-hand chain, it functions structurally as an orphaned second hemistich. This arithmetic subtraction creates a sudden architectural void in the stanza, introducing an entirely new rhythmic and visual dynamic without ever violating the foundational syllabic laws of *vina na mizani*.

The second innovation pushes the poem into the graphic dimension and gives the new form its very name: *Msisitizo* (emphasis). On the fifth line of each stanza, Mnyampala introduces a purely typographic structural marker to create this literal emphasis. The final word of the first hemistich—the exact word carrying the internal rhyme (such as *manono, patano, kinzano*, or *maridhiano*)—is visually emphasized by

being printed in boldface type. This typographical game cannot be sung; it is an operation of graphic reason. By visually isolating and stressing a structural metric pillar on the printed page, Mnyampala proves that the poetic grid is no longer merely a script submitted to a singer (*mghani*). It has become a sovereign visual object designed by an architect (*mtunzi*) for the eye of a reader, locking the whole system into strict syllabic compliance while demonstrating the infinite plasticity of its parameters.

5.6.3. The *Ngonjera* (The D*isputatio*)

Finally, Mnyampala transformed the classical meters into a dynamic, versified public dialogue with the *Ngonjera* (the *disputatio*). Derived from a *Cigogo* word meaning *ubishi* (argument or dispute) in Swahili, the *Ngonjera* functions as a poetic *disputatio*. The formal architecture of this meter is explicitly documented in the poem *Ngonjera ni kitu gani?* (What is Ngonjera?) (Mnyampala, 2013b, pp. xix-xxii).

The structural innovation of the *Ngonjera* lies in the introduction of a formal dialogic architecture directly into the strict *shairi* meter. The poem pits two named interlocutors against each other—such as *Mdadisi* (the inquirer) and *Jawabu* (the respondent)—who debate and contend stanza by stanza. Mnyampala popularized this versified *disputatio* to stage ideological debates during the Tanzanian *Ujamaa* era, making it a massively successful tool for political education. The form remains extremely popular in Tanzania today, even though its specific political origins and Mnyampala's role as its popularizer are often forgotten by the general public.

However, from a strict philological standpoint, Mnyampala's status as the absolute "inventor" of versified dialogue must be nuanced. The architectural integration of characters speaking and responding within metrical stanzas already existed in 19th-century Swahili literature, notably within the *utenzi* genre, as seen in classical epics like the *Utendi wa Haudaji* (Vierke, 2011). In this *utenzi,* characters' voices are continuously woven into the uninterrupted succession of stanzas—Mnyampala's generative leap consisted in transposing this dialogic mode into the *shairi* meter and radically altering its textual layout. In the *Ngonjera*, character names are extracted and positioned externally to the verses, physically breaking the standard stanzaic progression. This structural shift transforms the poem into a scripted, rhymed dialogue, effectively creating a formal hybrid between the classical *shairi* and the typographical conventions of European theater.

To modernize Swahili poetry, Mnyampala relied entirely on the immense structural flexibility already contained within the traditional metrical grid. The fact that the classical matrix can seamlessly accommodate a modern political *disputatio* perfectly validates the definitive observation made by the scholar Ibrahim Noor Shariff regarding the system's generative power: "[...] there is a great deal of freedom in composing with rhymes *[vina]* and measures *[mizani]*." (Shariff, 1988, p. 48)[61]

Ultimately, Mnyampala and modern innovators like Hashil prove that the strict rules of *vina na mizani* do not paralyze the poet. They offer a sovereign, mathematically precise foundation for infinite structural and ideological expansion.

Focus 1: Charles M. Mnyampala and the Preservation of the Archive

For over forty years after his father's death in 1969, the late Charles M. Mnyampala acted as the sole custodian of these unpublished archives. While the Tanzanian academic community was aware of the existence of these documents, locating them physically had proven elusive. In 1995, the Tanzanian scholar Professor Mugyabuso Mlinzi Mulokozi documented the missing link, noting: "Mary [Mathias Mnyampala's widow] believes that Mnyampala's unpublished manuscripts are with his son, Charles, whom we're still trying to contact" (Mulokozi & Sengo, 1995, p. 48).

In 2007, a geographical indication provided by Professor Mulokozi prompted our trip to Dodoma. Relying on serendipity and without guarantees of his presence, this journey led to a direct meeting with Charles, who was by then retired. This encounter established the connection that allowed for the complete on-site digitization of the archival corpus by the present author. Part of these previously unpublished manuscripts has since been published, and selected items from the archive are deposited in open access.

While academic researchers are dependent on Charles M. Mnyampala's preservation to access these primary sources, his contribution went beyond physical conservation. Alongside his professional career as a police officer (*mkuu wa polisi*) in the southern Tanzanian city of Mtwara, he was an independent researcher and a poet. His *dibaji* (foreword) to the posthumous publication of his

[61] Swahili original : "[...] upo uhuru mwingi katika kutunga kwa vina na mizani."

father's autobiography, *Maisha ni kugharimia* (Mnyampala, 2013b, pp. xi-xxvi), provides a structural synthesis of his father's compositional styles.

He systematically describes the organization of the corpus, detailing and naming the exact terms for specific rhyme patterns such as *mtiririko*, *msisitizo*, *ngonjera*, and *vidato*. By providing concrete examples for each named form, his *dibaji* serves as a practical introduction to Swahili poetic structures. His physical preservation of the archives and his intellectual insights into the texts are therefore inseparable and remain foundational to modern research on this corpus (see Chapter 8, section 8.5.4. on *Mashairi ya Vidato*).

5.7. Conclusion: Listening to Swahili Voices

This research does not claim to "discover" the *vina na mizani* system. Ascribing the posture of a discoverer to the modern researcher constitutes a profound misinterpretation of the empirical process. As this work demonstrates, the primary function of this chapter is to translate and foreground the meta poetical and theoretical discourse of the poets themselves. From Kaluta Amri Abedi to Shaaban Robert, from Ahmed Sheikh Nabhany to Ibrahim Noor Shariff, and from Hashil Seif Hashil to Mathias E. Mnyampala, this system has been explicitly theorized, formally codified, and either revered as dogma, expanded upon, or challenged as a strict boundary by generations of East African masters.

The primary sources compiled and analyzed throughout this chapter reveal a highly elaborate, endogenous metrical framework founded entirely on strict syllabic principles. Across the diverse theoretical writings of these practitioners, the concept of the accentual foot is entirely absent. The creators describe their generative engine with absolute clarity: the architectural foundation of classical Swahili poetry relies exclusively on the arithmetic precision of the syllable count (*mizani*) and the topological organization of the rhyme chains (*vina*).

To discard this established endogenous consensus in favor of imported Arabic or classical Greek metrical theories, and to artificially graft an invented accentual superstructure onto this strictly syllabic foundation, is more than a simple academic disagreement. It is a form of epistemic marginalization that prioritizes theoretical projection over the empirical reality of the creators themselves. The masters have spoken; it is time the academy listens.

Yet, the deafness of the accentual school is so entrenched that qualitative debate is no longer sufficient to dismantle it. To end this rigged analytical game, research must pivot to the empirical verification.

However, before deploying this empirical machinery, it is necessary to examine a striking historical paradox. As Chapter 6 will demonstrate, the accentual bias is not an inherent characteristic of all Western scholarship. Early European scholars, most notably Lyndon Harries and Jan Knappert, paradoxically operated as fiercely conservative custodians of the *vina na mizani* framework. They proved to be far more protective of the strict endogenous syllabic architecture than the subsequent wave of modern metrical theory.

Chapter 6. The Marginalized Legacy: How the Accentual School Overlooked the Syllabic Principles of Swahili Metrics in Jan Knappert and Lyndon Harries

Before concluding this series of chapters on the erasure of endogenous poetic sovereignty, it is intellectually imperative to clarify the boundaries of my critique and to expose a profound historiographical paradox. Rebutting the accentual school is by no means a blanket condemnation of Western academia as a whole.

To assume the existence of a coordinated ideological syndicate—as if a clandestine cabal of European philologists were secretly gathering in a basement to engineer the disqualification of African metrics—is a theoretical absurdity. There is no monolithic "Eurocentric" establishment acting in unison; such a broad categorization would be a failure of scientific precision and the expression of an essentialist ideology. My methodological objection is surgically targeted at a highly specific, localized lineage of thought: the accentual school of classical Swahili poetry.

6.1. The Francophone Epistemological Interlude: Charles Sacleux, Kris Berwouts and Contemporary Empirical Research

To fully grasp the magnitude of the empirical fidelity that exists outside this closed continuum, it is necessary to recognize that the faithful documentation of the endogenous system by exogenous observers extends far beyond the English-speaking or German-speaking academic spheres. The assumption of a monolithic, unified Eurocentric establishment imposing accentual frameworks is completely shattered when we examine Francophone research. This is a critical epistemological expansion, as a massive portion of the Swahili-speaking world—most notably the Democratic Republic of the

Congo—operates within the Francophone sphere, making French-language scholarship a highly legitimate, central, and indispensable component of any exhaustive bibliographical review. Within this Francophone tradition, however, it is Charles Sacleux who is consistently cited — unsurprisingly, since his conclusions serve the accentualist case — while syllabists such as Berwouts have received markedly less attention in the secondary literature. This imbalance reflects patterns of citation and canon formation, not the soundness of the underlying scholarship: writing decades after Sacleux, Berwouts had access to a far larger body of documentation than his predecessor ever did.

Father Charles Sacleux occupies a mixed position. He acknowledges the syllabic principle underlying the poets' practice while also introducing accent as a metrical principle in its own right, describing, in effect, a hybrid metrics—at once syllabic and accentual (Sacleux, 1909, pp. 8, 327). His grammar states plainly that Swahili versification rests on two axes: the syllabic count and accent together on one side, rhyme or assonance on the other, and he is careful to specify that the accent in question is not the tonic accent proper to each word in isolation, but what he calls the "accent oratoire [oratorical accent]" (Sacleux, 1909, p. 327)—a formulation of a concision that, to date, remains unmatched. The verses in common use run to 6, 8, 10, 12, 14, and 16 syllables, each with its own fixed caesura: after the fifth syllable in decasyllables, the sixth in dodecasyllables, the sixth (less often the fifth, eighth, or tenth) in fourteen-syllable verses, while sixteen-syllable verses form a doubled octosyllable closing on a single final rhyme. Each length carries, in turn, a fixed accent count — two in verses of three, four, six, and eight syllables, three in decasyllables, four in verses of twelve and fourteen — though Sacleux allows the syllable count itself to relax in longer verses, provided the rhythmic accents are observed; hiatus, too, is tolerated, if often absorbed by elision or contraction.

Sacleux defines the oratorical accent itself several pages earlier:

> Oratorical accent.— The tonic accent belongs above all to ordinary delivery and simple narration; in emotionally charged speech, it yields to the oratorical accent. By the latter is meant the accent that the speaker places on the words meant to convey the feeling he experiences. It is a free accent, both in the place it occupies and in the character it assumes. As to character, the syllable struck by the oratorical accent may add long or short

> duration to pitch height, or, in certain cases, cause intensity to predominate — with short duration, with or without pitch height. As to position, the oratorical accent either displaces the tonic accent, carrying it elsewhere, or uses it, merging with it on the same syllable. (Sacleux, 1909, p. 8)[62]

The definition is worth pausing on, because it sits uneasily with the fixed accent counts Sacleux assigns to each verse length. An accent that is "free" both in placement and character, and that belongs to the register of emotionally charged, declaimed speech, describes something closer to a performative phenomenon—realized in real-time when the verse is sung or chanted by the performer (*mghani*)—than to a structural feature permanently fixed at the moment of composition by the poet-architect (*mtunzi*). Sacleux does not rigorously distinguish between these two distinct domains, and the resulting ambiguity—between the invariant meter as structurally composed by the *mtunzi* and the fluctuating rhythm as acoustically performed by the *mghani*—is one this study will need to clarify. What his account offers regardless, and what precisely distinguishes it from most of what follows in this accentualist tradition, is clinical precision: a specific, countable claim for each verse length, open to empirical verification or refutation, rather than a mere impressionistic gloss.

While Sacleux's early twentieth-century account left a theoretical vulnerability regarding performative accents that later exogenous accentualists would recklessly exploit, this performative ambiguity is spectacularly absent in the subsequent rigorous scholarship of Kris Berwouts. This substantial study, published in the *Cahiers du CEDAF* under the title *Le sein de la mère: Introduction à la littérature classique et moderne en swahili* [The Mother's Breast: Introduction to Classical and Modern Swahili Literature].[63] (Berwouts, 1991) stands as a paramount instance of exogenous scholarship that

---

[62] French Original: "**Accent oratoire.** — L'accent tonique est surtout spécial au débit et à la simple narration; dans le discours ému, il cède le pas à l'accent oratoire. On entend par ce dernier l'accent que l'orateur met sur les mots destinés à traduire l'impression du sentiment qu'il ressent. C'est un accent libre, et par la place qu'il occupe, et par le caractère qu'il revêt. Quant au caractère, la syllabe frappée de l'accent oratoire peut ajouter la durée longue ou brève à la hauteur, ou, en certains cas, faire prédominer l'intensité avec durée brève avec ou sans la hauteur. Quant à la position, ou bien l'accent oratoire déplace l'accent tonique pour le transporter ailleurs, ou il l'utilise pour se confondre avec lui sur la même syllabe."

[63] The title winks at Shaaban Robert's well-known poem "Kiswahili: *titi la mama li tamu, hata likiwa la mbwa* [a mother's breast is sweet, even were it a dog's.]"

documents the endogenous Swahili metrical system on its own terms, without projecting any exogenous accentual scheme onto it. When Berwouts defines the architecture of the classical *utendi* (or *utenzi*), he relies strictly and exclusively on the precise arithmetic counting of syllables and the spatial alignment of rhymes. In his exhaustive description of the poetry, there is a thunderous and absolute absence of any notion of an accentual foot, a classical Greek trochee, an underlying temporal beat, or a phantom rhythmic pulse. He observes the texts with clinical precision, noting exactly how the architecture is constructed:

> The *tendi* are characterized by stanzas of four lines. Most *tendi* count eight syllables per line but several have six, nine, ten, eleven and even twelve syllables. The first three lines of each stanza end with the same syllable while the final syllable of the fourth line is maintained throughout the entire poem. (Berwouts, 1991, p. 26)[64]

This passage is significant for what it does not contain. Berwouts defines the meter through its stanzaic boundaries, its rhyming homophony, and the absolute syllable count of its lines, cataloguing the arithmetic variation of the *tendi* across poems from six to twelve syllables. At no point does the description require, or invoke, an accentual foot. It should be said at once that this catalog does not by itself refute an accentual model: Sacleux, precisely, assigns a different number of accents to each verse length, and any such model can scale with the line. What Berwouts establishes is something else, and it is enough: a Francophone scholar writing decades after Sacleux, with a far larger documentary base than his predecessor, describes the entire architecture of the *utendi* without ever needing the notion of a beat.

Furthermore, to bolster this severe formalism, Berwouts explicitly summons the voices of African scholars writing in English, creating a powerful intertextual consensus that completely bypasses the accentualist echo chamber. He quotes the Anglophone scholar Ireri Mbaabu to underscore the unyielding structural rigidity of the Swahili metrics.

---

[64] French Original: "Les tendi se caractérisent par des strophes de quatre lignes. La plupart des tendi comptent huit syllabes par vers mais plusieurs en ont six, neuf, dix, onze et même douze syllabes. Les trois premiers vers de chaque strophe se terminent par la même syllabe tandis que la syllabe finale du quatrième vers est maintenue à travers tout le poème."

> There are formal tunes of singing Kiswahili poetry. These tunes are governed by the number of syllables in a line and the rhythm and rhyming of such syllables. This is the primary reason that Kiswahili poetry is so systematic and rigid in its structure. It is due to such rigid insistence on poems which can be sung as Amri Abedi has stated, old pieces are preserved orally from generation to generation. (Mbaabu, 1985, as cited in Berwouts, 1991, p. 26)

Mbaabu's statement perfectly aligns with the endogenous reality: the poetry is "systematic and rigid in its structure" precisely because it is governed by "the number of syllables in a line" and the "rhyming of such syllables". His "rhythm" belongs to the tunes of singing he is describing, that is, to performance: what he says governs those tunes is the syllable count and the rhyme. The performative aspect ("formal tunes of singing") is presented as a direct consequence of this rigid syllabic foundation, not as an excuse to invent flexible accentual compensations or to project imaginary rhythmic stresses[65].

Furthermore, Berwouts' rigorous structural fidelity extends far beyond the *utenzi* meter; he applies the exact same arithmetic logic to the *shairi* and *wimbo* forms. When analyzing the *shairi*, Berwouts isolates the foundational elements of the stanza (*ubeti*), relying explicitly on the absolute counts of syllables (*mizani*) and lines (*mistari*). Citing the scholar Rajmund Ohly, he details the structural perfection of the form.

[65] A literal reading of Mbaabu's statement reveals a lingering terminological ambiguity: he does not explicitly differentiate between the structural composition meticulously engineered by the poet-architect (*mtunzi*) and the acoustic realization executed by the professional performer (*mghani*). However, an attentive reading of his text—without forcing the citation—reveals the implicit mechanics of this relationship. When Mbaabu speaks of a rigidity that governs the "tune," he is effectively acknowledging that the *mtunzi* provides a strictly predictable and unyielding syllabic matrix. It is precisely because this underlying structural foundation is so arithmetically stable and entirely devoid of unpredictable exogenous accents that the singer can safely trust the architecture, thereby gaining the absolute freedom to overlay their complex, variable vocal performance (*mahadhi*) onto the verse. Thus, while Mbaabu undeniably and correctly identifies the absolute purity of the syllabic system, his phrasing in this specific citation still slightly blurs the epistemological boundary between the invariant architectural blueprint and its subsequent musical execution.

> In its ideal form, a shairi presents itself, according to Ohly, as follows: each stanza (ubeti) consists of sixty-four syllables (mizani) divided into four lines (mistari) [...] There are therefore sixteen mizani in each mistari and each line is divided into two parts of eight syllables. (Berwouts, 1991, p. 42)66

It is essential to specify here, in light of the rigorous Swahili formal metrical theory, that the use of the term *mistari* (plural of *mstari*) designates the physical, graphic line on the printed page. As established in the broader theoretical framework of this manifesto, this graphic line is not necessarily aligned with the underlying structural line (*mshororo*), which is strictly delimited by the rhyming boundaries (*vina*). Furthermore, in the phrases "sixty-four syllables (*mizani*)" and " sixteen *mizani* " from this extract, the term *mizani* must be strictly understood as the "syllabic count" or "measure," rather than simply "syllable" (see Chapter 4 for a precise definition of the metrical terminology used in Swahili). Ohly and Berwouts' observation perfectly captures the romanized typographical layout of the verse, while fully acknowledging the underlying syllabic arithmetic that drives the true structural architecture. He then explicitly defines the architecture of the rhyme chains across the spatial boundary of the caesura:

> The first three half-lines preceding the caesura (*vipande vya kito*) share a common rhyme; the first three half-lines following the caesura also share a common rhyme, and the fourth half-line after the caesura features yet another rhyme, which is found throughout the entire poem. (Berwouts, 1991: 42)67

This description is of paramount historiographical importance because it directly anticipates the formal modeling that will be developed later in Chapter 8 of this manifesto. Although Berwouts expresses this architecture in prose and basic visual schematics rather than strict formal

66 French Original: "Dans sa forme idéale, un shairi se présente, d'après Ohly, comme suit: chaque strophe (ubeti) consiste en soixante-quatre syllabes (mizani) réparties en quatre vers (mistari) [...] Il y a donc seize mizani dans chaque mistari et chaque vers est divisé en deux parties de huit syllabes."

67 French Original: "Les trois premiers demi-vers qui précèdent la césure (vipande vya kito) ont une rime commune; les trois premiers demi-vers qui suivent la césure ont également une rime commune et le quatrième demi-vers après la césure présente encore une autre rime, que l'on retrouve dans tout le poème."

models, his conceptualization relies on the exact same underlying logic: spatial boundaries, precise syllabic counts (sixteen syllables divided strictly into two symmetrical blocks of eight), and the specific topological trajectories of the rhymes spanning across the hemistiches. In this strictly arithmetic definition, there is no theoretical void left for an accentual foot.

Likewise, when discussing the *wimbo* (plural *nyimbo*), Berwouts correctly identifies its deeply endogenous roots. He contrasts it with the Arabic etymology of the *shairi* by tracing the term *wimbo* directly to the Bantu verbal root *kuimba* ("to sing") (Berwouts, 1991, p. 55). While acknowledging the thematic porosity between *mashairi* and *nyimbo*, he observes the structural variations that occur in practice, noting that compositions can consist of stanzas of three or four lines with the number of syllables varying strictly between eight and sixteen per line (Berwouts, 1991, p. 43).

Whether analyzing an *utenzi*, a *shairi*, or a *wimbo*, the Francophone scholarship represented by Berwouts firmly anchors the analytical framework in the absolute measurement of the syllable and the spatial locking of the rhyme. This textual observation makes no use of the accents that some exogenous scholars desperately sought to impose.

Yet, remarkably, this substantial body of Francophone scholarship, is entirely absent from the bibliographies of the modern accentual school. Kris Berwouts' 1991 treatise is omitted both by Clarissa Vierke (2011) and Emiliano Minerba (2025), despite its detailed literary analysis of Swahili poetry and its exposition of the endogenous Swahili terminology governing poetic composition.

This historiographical blind spot applies equally to contemporary empirical research. Our own doctoral dissertation (Roy, 2013b), whose first published volume is devoted to metrics and bears the explicit subtitle *Définitions de la poésie d'expression swahilie* [Definitions of Kiswahili Poetry] (Roy, 2014, Vol. 1), is indexed and available in open access, and is not cited in Minerba's 2025 monograph on the same subject. The accentual paradigm does not even engage with this data-driven Francophone scholarship to refute it; it simply acts as if it does not exist, thereby immunizing its exogenous models from empirical friction.

The significance of these omissions extends well beyond individual bibliographical references. Berwouts' work and our own research are methodologically distinct and were conducted entirely independently of one another. Berwouts adopts a literary approach grounded in the Swahili conceptualization of poetry, whereas our own research is equally grounded in the Swahili conceptualization of the art while developing a

formal and quantitative analysis of its metrical structures. Nevertheless, both studies are data-driven, deriving their conclusions inductively from the observation of poetic corpora rather than from a priori theoretical assumptions. Although they employ different analytical frameworks, they independently describe Swahili metrics as a syllabic system and provide no empirical evidence in support of an accentual interpretation. This convergence is methodologically significant precisely because it emerges from independent lines of inquiry rather than from a shared theoretical framework or scholarly tradition.

From this perspective, the issue is methodological rather than personal. A comprehensive review of the literature on Swahili metrics would normally be expected to engage with empirical studies that directly address the same formal questions, regardless of their language of publication—assuming linguistic accessibility—institutional affiliation, participation in established academic networks, or the independent status of their authors. Yet the bibliographies of Vierke (2011) and Minerba (2025) largely overlook a substantial body of relevant Francophone scholarship on Swahili poetry. The result is not simply an incomplete bibliography but an incomplete representation of the available empirical evidence on the organization of Swahili verse.

This pattern raises a broader methodological concern. By overlooking methodologically independent studies that converge on the same empirical description of Swahili metrics, the accentual school risks restricting its evidential base to scholarship produced within a relatively closed citation network. Such a bibliographical configuration favors the internal reinforcement of existing theoretical assumptions while limiting systematic engagement with alternative empirical analyses. The consequence is not merely an incomplete survey of the literature but a narrowing of the empirical foundation upon which competing metrical models should be evaluated.

This phenomenon may appropriately be described as a form of intellectual feudalism: a scholarly ecosystem in which academic authority is transmitted primarily within established networks rather than continually tested against the broadest possible range of empirical evidence. Such a system does not necessarily exclude alternative research deliberately; rather, it tends to reproduce its own intellectual boundaries through ordinary bibliographical practice. If this dynamic could, at first glance, be attributed simply to the limited circulation of Francophone scholarship, a closer examination reveals a far more significant phenomenon. The same selective logic also affects the treatment of the foundational Anglophone literature that established the discipline itself.

It is therefore necessary to examine how the works of Jan Knappert and Lyndon Harries—two main architects of the modern transmission of Swahili poetry—have subsequently been received and reinterpreted.

While the accentual school might claim bibliographical ignorance regarding Francophone research, their treatment of foundational Anglophone scholars reveals a mechanism of selective amnesia. Contemporary accentualists do not explicitly deny the existence of the syllabic system documented by their predecessors; rather, they dismiss it as an obvious, uninteresting triviality. They bypass the observable arithmetic and syllabic reality in a speculative quest to uncover a supposedly "deeper" or "more fundamental" accentual logic—a hidden rhythmic pulse that, astonishingly, escaped the conscious awareness of the Swahili master poets themselves. To understand how this school managed to impose such an exogenous framework despite the overwhelming evidence, one must examine how Jan Knappert and Lyndon Harries—scholars who accurately documented the endogenous reality of the texts—have been paradoxically handled. The central paradox of Swahili metrical studies is this: the very ideological paradigms for which the most prolific Western transmitters of Swahili texts are condemned today—Orientalism and colonial classicism—were precisely the restrictive lenses that guaranteed their absolute fidelity to the formal mechanics of the *vina na mizani* system, the strictly syllabic framework that governs Swahili formal metrics exclusively through the exact enumeration of syllable counts (*mizani*) and the strict alignment of rhymes (*vina*).

6.2. Jan Knappert and the Preservation of the *Vina na Mizani*

To prove that Western academia is entirely capable of empirical fidelity to the Swahili poetical metrics, one need look no further than Jan Knappert. He embodied what the French term a *passeur*—an intellectual voyager and custodian of worlds who, despite his selection biases toward Islamic and didactic verse, actively rescued, preserved, and transmitted a fragile heritage across cultural boundaries. The postcolonial critique has denounced his Orientalism, and as scholars such as Alamin Mazrui and Ibrahim Noor Shariff have demonstrated, Knappert systematically over-emphasized the Perso-Arabic origins of Swahili culture, effectively disembedding the poetry from its African, Bantu epistemological

foundation. Mazrui and Shariff highlight how his scholarship operated under a pervasive "Arab-Islamic bias," noting that Knappert's views are typical of academic approaches that attempt to trace "each aspect of Swahili culture, literature and mythology back to its roots in one of the Oriental lands" (Mazrui & Shariff, 1994, p. 95). By doing so, he was explicitly "denying the Swahili their Africanity" (Mazrui & Shariff, 1994, p. 41). Mazrui and Shariff rigorously deconstruct this prejudice, arguing that such "modal and racial biases of Eurocentricity thus combined on the shores of Swahililand to 'conspire' against Swahili literature's secular verse in favor of its Islamic verse that supposedly betrays its Oriental connections" (Mazrui & Shariff, 1994, p. 96). They firmly state that a thorough survey of the tradition "reveals how mistaken Knappert's opinions are," as there is no point in their history where the Swahili produced a greater proportion of homiletic, religious verse than secular verse (Mazrui & Shariff, 1994, p. 96).

Yet, these same critics offer a vital, balanced concession that prevents a simplistic or purely vindictive dismissal of his legacy. Acknowledging his indispensable and pioneering role in the field, Mazrui and Shariff note that "Knappert's has been a distinguished career of research and writing on Swahili life and lore, and he is deservedly famous for assembling and disseminating a very considerable archive of Swahili poetry" (Mazrui & Shariff, 1994, p. 41). Consequently, they explicitly conclude that despite his ideological blind spots, "it would therefore be unfair to accuse him of anti-Swahili sentiments" (Mazrui & Shariff, 1994, p. 41).

Their critique also bears on selection: Knappert's collections tended to focus predominantly on religious themes (Mazrui & Shariff, 1994, p. 96). This editorial selection created a notable representational bias in Western-produced academic anthologies—shared to varying degrees by figures like Harries and J.W.T. Allen (Mazrui & Shariff, 1994, p. 96)—when contrasted with the vast corpus of secular editions published strictly in Swahili, whose organicity does not derive from a European reconstruction assembling various extracts from multiple authors, but rather from endogenous African literary production and curation—whether through single authorship or collections structured by local intellectuals—such as the *Diwani ya Muyaka* (compiled through complex cross-cultural collaborations) (Hichens, 1940) and the numerous collections published by local poets like Shaaban Robert, Amri Abedi, and Ahmed Nassir Bhalo (Mazrui & Shariff, 1994, p. 96).

Their concession, however, stands: his archiving career remains a cornerstone of textual preservation. For a data-driven metrical analysis, the primary imperative is the quantitative availability of reliable raw data; provided the internal architecture of the text has not been manipulated or distorted, every verified corpus remains analytically invaluable, regardless of the compiler's initial thematic preferences.

Operating within a Popperian epistemology, we sought to empirically verify the ideological critiques formulated by Mazrui and Shariff against the archives compiled by European scholars like Jan Knappert. It is undeniable that Knappert dedicated extensive, specialized monographs to Islamic verse, most notably *Traditional Swahili Poetry: An Investigation into the Concepts of East African Islam* (1967) and his monumental three-volume *Swahili Islamic Poetry* (1971). However, thematic specialization is a standard scholarly prerogative; the decision to compile exhaustive volumes on religious literature does not inherently equate to a deliberate erasure of the secular tradition. When subjecting the broader corpus to empirical scrutiny, the assertion that scholars like Jan Knappert and Lyndon Harries operated through an exclusively religious, Orientalist selection bias must be nuanced. A review of their general anthologies—specifically verifying the thematic diversity in Harries' *Swahili Poetry* (1962) and conducting a detailed textual analysis of Knappert's *Four Centuries of Swahili Verse: A Literary History and Anthology* (1979)—reveals a far more complex textual reality. These collections contain a substantial amount of secular material, including lyrical love poetry and historically grounded verses detailing the anti-colonial resistance against German forces (Knappert J., 1979, pp. 217-222). The thematic selection, therefore, is not nearly as absolute or exclusionary as the postcolonial critique implies.

Nevertheless, Mazrui and Shariff's methodological reproaches find partial validation when one examines the interpretive paratext, a bias we specifically verified in Knappert's work. Even when presenting secular or general poetry, Knappert tended to frame the Swahili poetic intent through a distinctly didactic, Islamic-adjacent lens. In his introduction to *Four Centuries of Swahili Verse*, under the heading "The Purpose of Poetry," he explicitly articulates this vision:

> The chief, if not the only end of poetry, Dryden said, is to delight. The Swahili people certainly take a great delight in their poetry; they speak of good verse as *mashairi matamu*, 'sweet poems', and they say:
>
> *Maneno yapendezayo* [*] *huleta mema matayo*
>
> Pleasing words, bring a good repute
>
> Yet neither fame nor the wish to please their fellow Swahili is the main inspiration for poets to write or recite their verses. Their chief aim is to teach. One poet, who admittedly belonged to the Islamic tradition, expressed it thus:
>
> *Kwa nufai siandiki*
> *mali ya dunya sitaki*
> *Nandika kwa Mwenye haki*
> *wana kuwapa wasia.*
>
> I do not write to gain advantage,
> the possessions of this world I do not crave;
> I write for the sake of Him who owns the truth
> and to admonish the young generation.
>
> The expression of these wishes is found throughout popular poetry in the world of Islam. Even many of the secular poets in East Africa write in a didactic spirit, so that a large part of Swahili poetry, Islamic and non-Islamic, narrative as well as non-narrative, can be called sententious or gnomic.
> (Knappert, 1979, p. xiv)

While this passage illustrates Knappert's epistemological tendency to subsume the entirety of Swahili poetics under a sententious framework rooted in the wider Islamic world, resolving this ideological debate falls outside the purview of the present study. We present these opposing arguments to acknowledge the complex historiography surrounding these texts, but as a data-driven scientist, my primary concern is strictly structural: is the data analytically sound? The fundamental question is not whether the compiler's interpretive paratext harbored an Orientalist bias, but whether the raw data—the endogenous architecture of the verses—remains uncorrupted. This is where the compiler's meta-vocabulary becomes essential. By explicitly recognizing and operating within the parameters of the *vina na mizani*—the traditional rhyming chains

(*vina*) and the strict syllabic count (*mizani*)—Knappert anchored his transcriptions in the authentic formal rules of the genre. As the following analysis will demonstrate, it is precisely this classical, conservative framework that paradoxically acted as the ultimate epistemological shield, ensuring the rigorous preservation of the Swahili metric system.

Therein lies the supreme irony of Knappert's legacy. Postcolonial and contemporary critics dismiss him for his perceived Orientalist biases, arguing that he viewed Swahili culture exclusively through a distorting Perso-Arabic lens. However, it was precisely this classical formation that compelled him to remain faithful to the poem's formal mechanics. Where ancient and modern accentualists see a rhythmic chaos that needs "fixing" through European stress-patterns, Knappert saw a rigorous, syllabic architecture. By treating the Swahili text with the same formalist deference one would accord to a Persian *qasida* or an Arabic *qasida*, he became the most reliable guardian of the *vina na mizani*—the very endogenous structure that defines the sovereignty of Swahili syllabic versification. His bias acted as an epistemological shield against the imposition of European accentual metrics. Across his monumental bibliography, his prosodic descriptions remain strictly empirical, respecting the architectural reality of the raw text.

6.2.1. The Empirical Evidence: The Absolute Syllabism in Knappert's Meta-Vocabulary

If the interpretive paratext of Jan Knappert leaned heavily toward an Arabo-Persian-centric and Islamic-focused worldview, his actual metrical descriptions of the raw data completely contradict the assumptions of the accentual school. A systematic review of his meta-vocabulary across his major works reveals that Knappert did not recognize, utilize, or superimpose any system of accentual feet or stresses. Instead, his analytical framework remained rigorously and exclusively syllabic, explicitly validating the endogenous *vina na mizani* system.

In his foundational work, *Traditional Swahili Poetry: An Investigation into the Concepts of East African Islam* (1967), Knappert establishes immediately that the defining characteristic of

Swahili poetics is its formal and structural rigidity, completely devoid of the ambiguities of stress-based systems. He observes:

> The term poetry, the definition of which presents considerable difficulty in a literature like English, can be easily defined in Swahili, where all poetry has a fixed metrical form, and is composed with very rigid patterns of rhyme (Knappert, 1967, p. 3).

When detailing the specific mechanics of the *utenzi* epic, Knappert leaves no room for accentual interpretation. The meter is defined purely by its strict syllabic count (*mizani*) and its rhyming chains (*vina*):

> Al[l] the poetic passages quoted in this work are composed in the *Utenzi* metre. There can be little doubt about the prosodic form of the *utenzi* metre: it always has eight syllables in the line and four lines in the stanza. The rhyme pattern is invariably a a a b or a a a a. The last rhyme is repeated throughout the poem in the last lines of all the stanzas. (Knappert, 1967, p. 5)

He further grounds this structural analysis in pure phonetics, removing any reliance on arbitrary stress, noting that:

> Uncertainty in the past over the number of syllables in the lines of Swahili poetry was caused by lack of knowledge of the prosodic value of the Swahili phonemes. A simple set of rules enables us to interpret the syllabicity of vowels and consonants. (Knappert, 1967, p. 5)

It is essential, however, operating within an epistemology of strict Popperian transparency, to address a singular, potentially confusing footnote in Knappert's 1967 text. When describing the *utenzi* stanza, he notes: "The rhythm is trochaic, identical with that of Longfellow's Hiawatha. See note 1 on p. 10" (Knappert, 1967, p. 5). The referenced footnote leads to an English translation by Mrs. J. H. Carter, who provided a "rendering in poetic English in the original rhythm," yielding lines such as: "Then a couch prepared she also / For the Prophet's own reposing" (Knappert, 1967, p. 10). To an English ear, this translation is indeed perfect trochaic tetrameter.

How do we reconcile this isolated reference to a "trochaic" rhythm with a strictly syllabic system? The answer lies in distinguishing between the structural rules of a meter and the phonetic epiphenomena of a language. In Swahili, natural word stress predominantly falls on the penultimate syllable. When a poet constructs a strictly measured eight-syllable line using Swahili vocabulary, the natural syntactic flow frequently produces alternating stresses. To a Western scholar educated in classical metrics, this *sounds* like a trochaic beat. However, this rhythmic oscillation is a phonetic byproduct of the language's morphology and syntax, not a generative metrical rule. In reality, as will be detailed through experimental verification in Chapter 9, the number of natural word stresses in an eight-syllable Swahili line is not fixed; it depends entirely on the natural accentuation of the words used. An 8-syllable line might contain two, three, or four stresses. Therefore, the systematic presence of the four stresses required to construct a true trochaic tetrameter is organically absent in the Swahili text. Knappert's reference to Longfellow's trochaic beat applies exclusively and descriptively to the *English translation* provided by Mrs. Carter (Knappert, 1967, p. 10). For the Swahili original, Knappert himself relies solely on syllabic principles to describe the poetry. This fact is proven by Knappert's own lexicographical work. In the dictionary co-authored with Leo van Kessel, he defines the *mizani* not as a metric foot, but strictly as the "minimum prosodic metrical element, syllable" (Knappert & Van Kessel, 2010, p. 432).

This purely syllabic paradigm is massively expanded and exhaustively documented in his later *magnum opus*, *Four Centuries of Swahili Verse: A Literary History and Anthology* (1979). Here, Knappert invalidates the premise that stress dictates form. Under the aptly named section *"The Metre of Swahili Verse,"* he formulates the absolute rule of Swahili metrics:

> Swahili meter is based on the number of syllables in the line. In Swahili verse, all the lines in one poem or song, have the same number of syllables (Knappert, 1979, p. 37).

There is no mention of rhythm or accentual feet determining the line length. The structural integrity relies entirely on the open Consonant-Vowel (CV) architecture of the Swahili language, which guarantees perfect syllabic isometry. Knappert catalogues an immense variety of metrical structures, defining each not by stress

patterns, but by the precise arithmetic of their hemistiches (*vipande*). He identifies dance songs and military verses by their exact syllabic counts, noting that even the most complex songs rely on this numerical regularity:

> The meter is 8+4 Syllables in the line in the following song [...] Another dance song from Tanganyika shows a similar structure (4+6 syllables) (Knappert, 1979, p. 37)
>
> Even a dance song which the Chaga in Tanzania sing in Swahili appears to have a regular form (4+4 syllables in the line) (Knappert, 1979, p. 38).
>
> Mnyampala gives 84 Beni dance songs (in Swahili, 1963), all in regular meters. The following song has the commonest song meters 6+ 6 syllables in the line [...] (Knappert, 1979: 38).

Knappert's documentation further details how lines exceeding eight syllables are systematically divided by a caesura into *kipande*:

> Most song stanzas in Swahili have three lines. If a line is longer than eight syllables, it has a caesura. Each 'piece' of a line is called a *kipande*. (Knappert, 1979, pp. 38-39)

He cites specific examples, such as wedding songs with "8+6 syllables" (Knappert, 1979, p. 39) and notes that "The 8-syllable lines of the following song are divided each into two *kipandes* of 2+ 6 syllables." (Knappert, 1979, p. 39)

He even highlights highly complex, asymmetrical structures that are incompatible with any theory of binary accentual feet. When discussing intricate pun-songs, he provides an example divided into three distinct hemistiches per line: "metre: 6+4+5, *ukawafi"* (Knappert, 1979, p. 42). The existence and documentation of a 6+4+5 meter stands as definitive empirical proof that the meter cannot be governed by stress; such an arithmetically uneven line can only be constructed, memorized, and sustained through strict syllable

counting. In historical contexts, he records the deeply ingrained nature of syllable-counting, noting that

> The couplet, Arabic *bayt*, Swahili *baiti* or *beti*, was once a pure Swahili word for lines of verse in general, and for the oldest preserved form of Swahili poetry, the distichs of 15-syllable lines, the only extant non-narrative meter before the end of the eighteenth century [...] *ukawafi*, usually meaning the specific term for the 15-syllable meter [...] (Knappert, 1979, p. 35)

Coupled with this syllabic precision is the rigid architecture of the rhyme (sg. *kina,* pl. *vina*). Knappert delves into its etymology and structural function:

> The Swahili word for rhyme is *kina*, which makes the plural *vina* or, in Siu, *zina*, which could be from the Arabic *mazini*, meter, but is more likely to derive from Swahili *jina*, 'name'. (Knappert, 1979, p. 36).

He observes the absolute phonetic matching required:

> In the oldest Swahili poetry we find the following rules of rhyme being applied. Rhyme is the identity of the last syllable. Syllables are (firstly) simple vowels, so that in the Herekali (1728), *ia* rhymes with *aa*, the final *a* being the rhyme. (Knappert, 1979, p. 36)

Furthermore, he emphasizes that:

> In Swahili poetry, two lines rhyme if their last syllables are identical; the vowel preceding the last syllable does not have to be the same. (Knappert, 1967, p. 5)

He provides the clinical example that in Swahili, rhyming sounds like *"kilo, silo or tango, bingo"* (Knappert, 1967, p. 5).

The rhyme is a morphological anchor, not an accentual cadence. He notes the historical evolution of this system, stating that:

> Whereas Arabic rhyme schemes have been fixed ever since the tenth century, Swahili poets are constantly developing new patterns, and weaving ever more rhyme into their lines. (Knappert, 1979, p. 36)

Furthermore, he emphasizes the inherently strophic nature of this poetry to contrast it with foreign models:

> Nothing prevents a poet from extending a three-line stanza into a long poem composed of several stanzas. All Swahili poetry is *muwashaha*, 'strophic' that is stanzaic; this in contrast to Arabic poetry which is only exceptionally composed in stanzas, mostly in single lines rhyming throughout the poem. (Knappert, 1979, p. 40)

He demonstrates how hemistiches and lines interlock through these end-syllables, noting cases where:

> For the following verse structure the poet has taken three *kipandes* of eight syllables together to form a stanza, a very old type of poem. (Knappert, 1979, p. 40)

In summary, when we strip away the interpretive, Orientalist paratext regarding the didactic or Islamic origins of the verse, the raw metrical data provided by Knappert is clinically and massively syllabic. Over the course of numerous pages, from the 8-syllable *utenzi* to the 6+4+5 *ukawafi* and the 15-syllable distichs, Knappert explicitly uses the *mizani* (syllable count) and the *vina* (rhyming syllable) as his exclusive descriptive tools. For the data-driven scientist, this meta-vocabulary constitutes an undeniable empirical record: the internal mechanics of classical Swahili poetry function independently of Western accentual theories.

6.3. Lyndon Harries and the Syllabic Architecture of Classical Swahili Poetry

Lyndon Harries shares a similar, albeit distinct, historiographical fate. Functioning within the colonial administrative apparatus, Harries produced scholarship that is simultaneously foundational and highly contested. It is essential to acknowledge that his monumental contributions are deeply understood and highly respected by East African, and specifically Tanzanian, researchers. As M.M. Mulokozi and T.S.Y. Sengo explicitly state in their *History of Kiswahili Poetry* (1995), his seminal 1962 publication remains an inescapable pillar of the field. They recognize that "the most comprehensive works on the

history of Kiswahili poetry so far are Swahili Poetry by Lyndon Harries [...] " (Mulokozi & Sengo, 1995, p. 3). This acknowledgment proves that his philological labor is not merely an exogenous relic to be easily discarded, but a cornerstone that continues to anchor modern academic discourse.

However, this profound scholarly respect does not shield his approach from severe methodological criticism regarding his restrictive definition of the genre. Harries has been rightly criticized for his literary paternalism and for attempting to freeze Swahili poetry within a rigid, normative classicism. As Mulokozi and Sengo point out, Harries largely posited that "authentic" Swahili literature had to be inherently antique and strictly Islamic to possess any legitimate value, thereby marginalizing contemporary, secular, or purely oral endogenous traditions. They underscore this severe ideological limitation by noting that "Lyndon Harries could still speak of Kiswahili poetry as an exclusively Islamic artform" (Mulokozi & Sengo, 1995, p. 21).

Yet, much like Knappert, his ideological flaws paradoxically ensured his absolute metrical accuracy. It was precisely this conservative obsession with classical purity and textual orthodoxy that drove Harries to document and preserve the uncompromising compositional rules exactly as they were transmitted. His conservatism did not lead him to artificially impose an alien or rigid accentual framework onto the verse; rather, it engendered a profound deference to the explicit theoretical framework articulated by the Swahili composers themselves. Because he valued the classical text above all else, he rigorously recorded the endogenous structural mechanics—the strict syllabic measurements and the spatial alignment of rhymes—and he preserved the exact Swahili metrical meta-vocabulary employed by the poet-architects (*watunzi*). By faithfully transcribing the creators' own definitions of their craft, Harries cemented the undeniable empirical reality: the endogenous system is exclusively and entirely defined by a syllabic *vina na mizani* metrics, leaving no theoretical void for exogenous accentual speculations.

6.3.1. Lyndon Harries and the Recognition of *Sui Generis* Syllabism

The empirical observations validating the syllabic nature of classical Swahili verse are powerfully articulated by Lyndon Harries in his seminal work, *Swahili Poetry* (1962). In the context of East African historiography, Orientalism typically manifests as an epistemological bias that instinctively locates the aesthetic, structural, and cultural genesis of Swahili artifacts within an Arabo-Persian and Islamic framework, often minimizing or erasing the endogenous African matrix. Lyndon Harries, as a scholar deeply immersed in Arabic traditions, indubitably approached Swahili literature bearing the hallmarks of this selection bias. He actively sought the Arabian background within the manuscripts. Yet, the profound significance of his analysis lies precisely in the fact that, despite this Orientalist lens, the empirical reality of the textual data forced him to acknowledge the absolute structural autonomy of Swahili verse. He concedes that while the thematic material may look toward the Orient, the poetic architecture is African.

Harries begins his study by firmly establishing this unique, endogenous nature, explicitly warning against the tendency to view Swahili poetry merely as a derivative outpost of Arabic models. He asserts with striking clarity:

> Although Swahili poetry derives from Arabic poetry, in many respects there are fundamental differences. It cannot properly be compared with Arabic poetry, for it is sui generis. The Arabist may recognize the obvious connexion between Swahili poetry and Arabic poetry, especially in much of its subject-matter, but Swahili poetry cannot be considered as an extension of Arabic poetry or even as a modification of it. Swahili poetry exists in its own right and therefore must be estimated, as to its merit, for what it is in itself. (Harries, 1962, p. 3)

This acknowledgment of the sui generis nature of the poetry represents a key epistemological step. Despite his bias, Harries realizes that the mechanics of the poetry defy Arabic categorization. He then nuances his position by recognizing that an understanding of both the Oriental and the African foundations is indispensable, creating a unique literary amalgam where the African structure

incorporates the Islamic worldview. He notes that the depth of the poetry cannot be grasped without acknowledging this dual heritage:

> At the same time, those Arabists who have had some experience of Swahili poetry do not fail to show the deepest interest and recognize that without a knowledge of its Arabian background even Swahili-speaking Africans may fail to understand that in their poetry the Swahilis have a possession of which they may well be proud. Similarly, without some experience of the African background of Swahili poetry the Arabist may not appreciate all that the poetry has to offer. From the amalgam of two traditions we find in Swahili poetry a positive contribution to literature. (Harries, 1962, p. 3)

Having established this autonomous Swahili literary space, Harries proceeds to detail the structural foundation of Swahili prosody, emphasizing the absolute primacy of the rhyme (*vina*) and the syllabic count (*mizani*). He observes that the verse (*baiti*) is defined by its rhyming system rather than by graphic layout:

> In Swahili a verse, *baiti* (Ar. *bait*), is often written as a single line, but there are many examples in which the verse is written in two lines with a medial caesura in each line. It is the rhyming system that provides the clue to the verse-form in Swahili poetry. (Harries, 1962, p. 9).

He elaborates on this internal architecture, showing how the rhyme dictates the physical division of the line:

> As in Arabic, the end of a verse is marked by a rhyme which is the same for all verses throughout the poem. Internal rhymes mark the ending of each section (Sw. *kipande*) of a verse. (Harries, 1962, p. 10)

Harries warns that Western translations and transcriptions often distort this reality by imposing visual line breaks that do not exist in the original manuscript tradition, reiterating that the structure is bound by the formal rules of the rhyme:

> In translating the poems into English it is customary to divide the poem into lines which, in fact, are not written as separate lines in the Swahili-Arabic script. This is done most often with verses

> rhyming abababbc, giving in the translation four lines, as follows ab ab ab bc when in the script the whole verse is written as a single line or as two lines with medial caesurae (Sw. *kituo/zituo*). Classification of Swahili poetry should be by the number of sections (Sw. *vipande*) in a verse and this is determined by the rhyming system. (Harries, 1962, p. 10)

However, it is in his clinical description of the meter itself that Harries delivers the most definitive and devastating refutation of the accentual school. He states that Swahili meter is solely a measure of counted syllables, explicitly and repeatedly denying any structural role for metrical feet or an overarching metric system of stress. This passage effectively dismantles any attempt to superimpose exogenous accentual frameworks onto classical Swahili poetics:

> Swahili metres are the measure of a section (Sw. *kipande*) in terms of the number of vocalized syllables (Sw. *herufi*, Ar. pl. *ḥarūf*) to the *kipande*. Vowel length is not significant, so that it is not possible, as in Arabic, to arrive at any system whereby the measure of syllables can take into account the metrical feet within a measure. In measures of Swahili verse there may be major (usually penultimate) and minor stresses, but there is no metric system of stress, nor can this be related to the subject of vowel-length or of metrical feet. (Harries, 1962, p. 10)

This statement is profound in its analytical precision. Harries perfectly acknowledges the phonetic reality of the language—noting the organic presence of natural "major (usually penultimate) and minor stresses"—but he denies that these natural speech accents form a "metric system of stress". The metric foundation is purely arithmetic and governed by formal composition, based entirely on the counted *mizani*. He confirms this endogenous terminology with absolute finality:

> The term applied to the number of vocalized syllables in a *kipande* is *mizani* (cf. Ar. *al-mīzān*). Verse of eight *mizani*, for example, has eight vocalized syllables to each *kipande* in the verse. (Harries, 1962, p. 10)

Finally, Harries demonstrates that any deviation from this strict dual requirement of exact syllabic count and precise rhyming is immediately classified as a structural failure (*guni*) by the Swahili

poets themselves. If stress were the organizing principle, a missing syllable could be compensated for by a heavy beat; however, in Swahili, a missing syllable breaks the verse entirely. He concludes his prosodic definition by highlighting this rigid structural law:

> A rhyme (Sw. *kina*, pl. *vina*) is monosyllabic in Swahili poetry and consists of consonant plus vowel. Verses defective in rhyme or in syllabic measure are called *guni*. Verses with defective rhyme are also called *mashairi ya kutupa vina*, i.e. verses discarding rhyme. (Harries, 1962, p. 10).

In conclusion, Lyndon Harries, a scholar deeply familiar with the Arabic roots of Swahili culture, nevertheless recognized the absolute syllabic nature of its poetry. His explicit rejection of metrical feet and metric stress, combined with his definition of the meter solely in terms of "the number of vocalized syllables" (*mizani*), provides irrefutable, independent verification of the endogenous structural rules governing classical Swahili verse.

By dismissing Harries entirely as a mere relic of the colonial era, modern scholars conveniently discard the empirical evidence he provided—evidence that directly contradicts the speculative frameworks of the accentual school. In this sense, Harries is not just a "colonial scholar" to be deconstructed; he is a witness whose technical accuracy has become a severe embarrassment to the accentual orthodoxy.

6.4. The Mechanics of a Scholarly Marginalization of Vina na M*izani*

Before turning to the mechanisms of this marginalization, the evidence gathered above should be weighed, since the three witnesses who converge on the syllabic account do not carry equal authority and it is better to say so than to let the reader assume a uniform one. Harries denies the existence of a metrical system of stress in explicit terms, and he does so as an Arabist who knew perfectly well what a metrical foot is; that is a positive statement about the object. Knappert defines the *utenzi* exclusively by syllable count and rhyme, and glosses *mizani* in his own dictionary as the syllable itself; that is a strong indication drawn from a working meta-vocabulary. Berwouts does not

mention accent at all; that is an argument from silence, and a weaker one, since he is writing a literary introduction and not a study of the meter, and his silence may reflect the scope of his book as much as a judgment about the verse. The convergence of the three is meaningful precisely because it was reached independently, but the demonstrative burden rests on the first.

What term, then, should we use to describe a situation in which the Swahili poets themselves explicitly define their own system in syllabic terms—*vina* and *mizani*—where the most prolific Western transmitters and philologists of their texts confirm this empirical reality through decades of textual work, and yet an entire academic lineage proceeds as though this convergence of endogenous and exogenous testimony barely existed?

Vierke (2011) and Minerba (2025) represent a recent phase of accentual scholarship. Their work will serve here as a point of entry for examining this central question. "Erasure" is too strong a word, and inaccurate: neither Vierke nor Minerba erases Knappert or Harries from their bibliographies. Both name them, repeatedly. What is at stake is not deletion but marginalization—a partial, structural amnesia that operates by displacement rather than suppression: the endogenous voice, and the philological witnesses who amplified it, are acknowledged as a formality, cited for peripheral or unrelated claims, and then swiftly bypassed the moment the analysis turns to what actually interests the accentualist school—the search for stress, beats, and feet. This pattern is not unique to one author; it is the shared operating procedure of the contemporary accentualist lineage as a whole, and examining Vierke (2011) and Minerba (2025) together, rather than separately, is what makes the mechanism visible as a structural habit rather than an individual lapse.

Begin with what both scholars themselves establish as ground truth, because their own formulations show how solid, how uncontested, and how quickly abandoned as an object of theoretical inquiry the syllabic system is in their accounts.

Minerba writes that "*mizani* is a concept equivalent to that of syllable, as explicitly stated by Amri Abedi (1954: 1)" and that "'Isosyllabism' is expressed in Swahili as *urari wa mizani* ('equality of the measures')" (Minerba, 2025, p. 156). He adds, on the same page, that "the word for 'rhyme' is instead *kina*" and specifies precisely how Swahili *kina* differs from rhyme in European traditions: "it consists, in fact, in the identity of the last syllable of two words, so that *nyuma* and *neema* form a *kina*, even if this would not be considered a rhyme

in other traditions" (Minerba, 2025, p. 156). Summarizing the whole classical system, Minerba states that "all the Swahili metres that exist in classical literature are characterised by three elements: the constancy of the number of syllables in all the lines (isosyllabism); the presence of stanzaic patterns based on regular distributions of rhyme; caesura" (Minerba, 2025, p. 156).

Vierke performs an almost identical gesture. *Mizani* and *vina* are named, in her own text, as "the two parameters which have frequently been pointed out in the literature on *tendi* as the most essential prosodic characteristics" (Vierke, 2011, p. 28), and she records her own informants Ahmad Nabhany and Ahmad Nassir insisting that "the line meter is primarily based on syllable count" and that "any deviation from the prescribed eight-syllable count is a fault (*guni*)" (Vierke, 2011, p. 28). Together, these two passages amount to a complete and undisputed technical restatement of what Knappert and Harries had already established philologically decades earlier: *vina na mizani* as the two constitutive parameters of the Swahili line.

What is decisive is what happens immediately after each of them lays down this definition—and it is here that the two authors converge most tellingly on a shared reflex. Minerba does not contest, qualify, or nuance isosyllabism as a raw textual datum; he states it, cites the emic terminology correctly, and then treats the matter as settled, a preliminary and self-evident fact of description rather than a subject demanding sustained theoretical elaboration, before moving on within the same section to caesura placement, stanzaic structure, and comparison of the *utenzi* meter with Arabic *tasmīṭ murabbaʿ*, explicitly leaning on Vierke to authorize this pivot:

> The pattern of *utenzi* has a structure which reveals great similarities with the Arabic *tasmīṭ murabbaʿ*, to the point that a derivation of the former from the latter is highly plausible, as remarked also by Vierke (2011: 51). (Minerba, 2025, p. 158)

In doing so, Minerba perfectly mirrors the methodological reflex of Vierke herself. Having named *mizani* and *vina* as the two essential parameters and having recorded her informants' insistence on the eight-syllable rule as an absolute law, she immediately concludes that "much less attention has been paid to rhythmicisation through accent or beat" (Vierke, 2011, p. 28)—as though this imbalance in the scholarly literature were simply a natural state of affairs to be inherited and continued, rather than a choice she is herself about to reproduce.

In both cases, the syllabic count becomes a mechanical baseline, a piece of technical housekeeping dispatched in a single subsection before the real analytical apparatus—the rhythmic core, the beat, the *Grenzschema*—is deployed as the properly theoretical object of inquiry.

The epistemological necessity of this marginalization becomes overwhelmingly clear when we dissect the internal fragility of the models they subsequently attempt to impose. Engaging directly and deeply with the strictly arithmetical and syllabic principles transmitted by Knappert and Harries would instantly falsify the exogenous accentual paradigm. Therefore, they must marginalize this formal reality to protect their own theoretical epicycles.

For Vierke, this means systematically sidestepping the endogenous formal syllabic grid of Swahili poetry to justify her *Grenzschema*—an idealized 4+4 cognitive prototype (see Section 2.3.1). As Knappert and Harries rigorously documented, the endogenous principle of *mizani* defines the segments (*vipande*) exclusively by their absolute syllabic capacity. When a composer fills a fixed eight-syllable *kipande* with Swahili vocabulary, the strict penultimate stress rule of the language deterministically assigns the accents and phonological word boundaries. Depending entirely on the morphological size of the words selected to fulfill this exact syllable count, the eight-syllable space will statistically yield one, two, three, or even four distinct lexical stresses as seen in this octosyllabic example *ḿtu húyu ána mámbo* ("This person has issues"). Because this deterministic phonological reality routinely shatters her required two-beat symmetry, Vierke is forced to execute a profound theoretical contortion. She must abandon the accentual structure that the words of the line actually deliver, and engage in blatant cherry-picking from an undocumented "chanted" performative mode (Layer 3). By selectively isolating performative anomalies—such as her specific reproduction of stanza DA 159 (*kwa mámbo bora adhímu*) (Vierke, 2011, p. 105), where the realization she transcribes, most likely a chanted one, leaves the lexical stress on the adjective *bora* unrealized to conveniently leave exactly two beats (*mámbo… adhímu*)—she forces the data to fit her exogenous model. Furthermore, she shields this selective sampling through a practice of archival gatekeeping. By omitting the comprehensive accentual marks from Dammann's manuscript in her published edition, she effectively prevents the reader from accessing the complete historical dataset.

For Minerba, the marginalization of the endogenous syllabic matrix is required to sustain his diffusionist hypothesis. If Minerba were to fully acknowledge the conclusions of Harries—that Swahili poetry is sui generis and that its internal engine was stripped of Arabic metrics the moment it adopted the Bantu phonological reality—his entire theory of the "creolized trochee" would collapse. Instead, Minerba falls victim to a homonymic illusion: he mistakes the spatial borrowing of a stanzaic shell (like the Yemeni *mubayyat*) for the importation of a deep metrical engine. To force this Arabic tetrasyllabic rhythm onto Swahili verse, he commits a textbook Texas Sharpshooter fallacy. He draws an artificial "trochaic" boundary around the natural, passive penultimate stress of the Swahili lexicon, falsely claiming a generative metrical intent where there is only an inescapable phonological reflex.

And here lies the crowning absurdity of the whole accentualist enterprise: the scholar sits down opposite the poet, listens to the poet declare in his own language that the octosyllable is law and any breach a fault, and then walks away convinced that he, the scholar, has uncovered the poem's true, deeper structure—a structure the poet supposedly built without knowing it, and can neither name in Kiswahili nor recognize even when it is pointed out to him. It is one thing to disagree with a native theorist on empirical grounds; it is another to declare oneself the custodian of an unconscious design that eluded its own architect. What kind of scholarship grants itself the authority to know a poetic tradition's foundations better than the tradition's own most competent practitioners? And, just as alarmingly, what kind of scholarship systematically marginalizes its own faithful exogenous transmitters—those conservative Western philologists who preserved the uncompromising *vina na mizani* framework in the empirical record—simply because their strict formal accuracy contradicts the new accentual dogma? By treating both the explicit, normatively enforced testimony of the endogenous creators and the rigorous textual documentation of early compilers as the naive surface of a truth only the modern analyst is equipped to perceive, this approach executes a profound double disqualification.

The truly striking feature is what Knappert and Harries are cited for, in both authors, and what they are never cited for. Minerba never invokes either scholar in his discussion of *mizani*, *kina*, or isosyllabism (Minerba, 2025, pp. 156, 158)—the very domain in which their philological authority would be most directly relevant. Instead, Knappert is cited for the sweeping cultural statement that "Swahili

culture is essentially Oriental, not African, in its material as well in its spiritual aspects" (Knappert, 1967, p. 9, as cited in Minerba, 2025, p. 13). A claim Minerba goes on to dismantle at length using Pearson and Spear; and for the unsupported genealogical claim that the Swahili quatrain *shairi* derives from the Persian *rubāʿī* rather than from Arabic metrics. Minerba explicitly criticizes this assertion, taking the opportunity to launch a severe ideological attack on Knappert's supposed denial of Swahili agency:

> One supporter of the first view was Knappert (Knappert, 1979: 59), who once remarked that the Swahili quatrain (*shairi*) has no counterpart in Arabic metrics; he concludes, presenting no supporting evidence, that its most probable ancestor was the Persian *rubāʿī* (whose structure, however, is very different from that of the *shairi*); one might wonder if, in Knappert's opinion, the Swahilis could ever have 'discover[ed]' the aesthetic effect of metrical patterns in poetry on their own. (Minerba, 2025, p. 5)

This rhetorical question stands as a monument of unintentional, twofold irony. First, Minerba explicitly dismisses Knappert for "presenting no supporting evidence" regarding his genealogical claims. Yet, Minerba's own foundational metrical theory—the "trochaic innovation" (Minerba, 2025, pp. 217-221) purportedly born from an Arabic tetrasyllabic template—is an entirely speculative construct devoid of any empirical proof within the Swahili lexicon (see Chapter 4) or the statistical reality of the classical texts (see Chapter 9). He attacks his predecessor for a lack of evidence while simultaneously building his own paradigm on an abstraction that, as he formulates it, is shielded from falsification. Second, Minerba postures as the defender of African intellectual sovereignty against Knappert's Orientalist prejudice. Yet, it is Minerba himself who systematically marginalizes the endogenous *vina na mizani* framework—the very formal system the Swahili poets *did* discover, codify, and enforce on their own, and which Knappert preserved in his empirical transcriptions. By dismissing the explicit, arithmetical and syllabic rules formulated by the Swahili master poets in order to force their verse into an imported, Arab-derived rhythm, Minerba actively commits the exact epistemological erasure he so piously condemns.

Harries fares no better: he is cited for the ethnographic-cultural assertion that the Swahili "may have Bantu blood, but their whole way of life is Muslim-Arabic" (Harries, as cited in Minerba, 2025, p. 17)—

again a claim about cultural origin, not about metrical structure, and one Minerba explicitly rejects as unsupported by the archaeological and commercial evidence he marshals from Pearson and Vernet. In other words, Knappert and Harries are cited by Minerba only in their capacity as commentators on cultural and historical origins, precisely the register in which Minerba wants to disagree with them, and never once in their capacity as the philologists who transcribed, edited, and annotated the very texts whose syllabic regularity Minerba himself independently confirms a few dozen pages later. The same asymmetry structures Vierke's engagement, inflected differently but following the identical logic. Harries and Knappert figure in her bibliographic apparatus and are invoked for matters of textual transmission, corpus availability, and editorial precedent—yet their descriptive claims about the primacy of syllable count and rhyme are never treated by her as a serious theoretical alternative to be argued against on their own terms.

What emerges from placing Minerba and Vierke side by side is therefore not a simple repetition of one scholar's individual oversight by another, but a shared structural habit across the new accentualist lineage as a whole: cite the philological authorities where their historical and cultural claims can be corrected or contextualized; never cite them where their observations on *vina na mizani* would require the accentual project to justify its own priority over the syllabic one. This is not falsification, and it is not a simple oversight born of unfamiliarity with the sources—both authors demonstrably know Knappert and Harries well enough to quote them precisely and to criticize them by name on other points. It is, rather, a selective citational practice that grants philological credit on secondary and correctable matters while withholding it on the one matter—the empirical, syllabic reality of *vina na mizani*—where crediting it fully would undercut the accentual paradigm's own claim to superior theoretical depth. Stripped of the exogenous accentual apparatus—the rhythmic core, the beat, the *Grenzschema*—the Swahili emic vocabulary itself, as both Minerba's technical exposition and Vierke's own informants describe it, names a self-sufficient, empirically verifiable, strictly syllabic architecture: *mizani* as the unit of measure of the line, *kina* as the governing principle of rhyme placement, and *urari wa mizani*, the equality of measures, as the constitutive law of the line. This system requires no accentual scaffolding to be legitimate as a metrical theory in its own right, and its formal sovereignty—established by the poets, restated without contradiction by their most

careful transmitters, and confirmed even by the accentualist scholars themselves the moment they turn to description rather than theory—stands unassailable.

Yet, as this historiographical review closes, a word must be said on what this chapter does and does not establish. It is a chapter about what observers of the texts have written, and testimony is not measurement. Nothing in it proves physically that Swahili meter is syllabic; that burden falls entirely on Chapter 9, where the models are confronted with datasets stated in advance and published in full. What this chapter establishes is narrower but still considerable: that across three independent scholarly traditions, and across sixty years, no observer working directly on the manuscripts found it necessary to invoke an accentual foot in order to describe the meter, and that one of them, having every reason to look for feet, explicitly recorded their absence. The accentual school has never refuted these descriptions. It has simply ceased to cite them.

Chapter 7. The Cost of Formal and Scientific Methods: Epistemological Rupture and Academic Inertia

The silencing of sovereign Swahili voices in the study of Swahili poetry does not end with the relegation of classical master poets to mere unconscious executors of an imported, phantom accentual pattern. This insidious process of invisibilization extends, by a relentless and predictable institutional logic, to any scholarly endeavor that undertakes to demonstrate the empirical validity of the endogenous system. It is important to clarify that this academic feudality does not represent an absolute, monolithic dominant paradigm, but rather a specific, dogmatic trend—the accentual school and its various avatars—which operates on rigid a priori assumptions, forcing texts into the Procrustean bed of exogenous metrics.

In its defense of the speculative framework of accentual metrics, the accentual school systematically marginalizes Western scholars such as Lyndon Harries and Jan Knappert. It is imperative to state our own position regarding these pioneers: We hold a profound respect for their contributions. Without their lifelong dedication to rescuing, preserving, translating, and editing the classical Swahili poetic corpus, a significant part of these historical texts would simply not be available to the modern world. The sheer scale and foundational value of their philological labor provide a massive, invaluable database—a quantitative reality explicitly acknowledged even by their postcolonial critics, such as Mazrui and Shariff (1994, p. 41), as well as Mulokozi and Sengo (1995, p. 3). For a data-driven scientist, this immense archival contribution outweighs the ideological critiques often leveled against them, provided, of course, that the raw data itself has not been subjected to any form of textual corruption or methodological distortion during the editing process.

These scholars have indeed been heavily criticized and explicitly accused of Orientalism—charged with applying an ideological filter that elevated the Islamic and Arabo-Persian elements of Swahili culture while allegedly minimizing its fundamental African roots. Yet, setting aside any speculation about their personal motivations or

academic ideals, this selection bias did not compromise the validity of their formal analyses. The key point here is not merely that the text itself remained uncorrupted, but rather that the analytical metatext they produced to describe these poems remained rigorously faithful to the endogenous model utilized by the Swahili poets themselves. By explicitly recognizing and operating within the parameters of the *vina na mizani*—the traditional rhyming chains (*vina*) and the strict syllabic count (*mizani*)—their descriptive framework effectively conserved the authentic structural rules of classical Swahili poetry. They documented the *vina na mizani* without ever feeling the theoretical need to superimpose an exogenous, artificial accentual metric onto the verses. Consequently, in a striking epistemological paradox, it is precisely this accusation of Orientalism that makes their metrical observations so particularly powerful today.

One would logically expect scholars operating under such a pronounced Orientalist bias to eagerly import Oriental metrical models—such as the quantitative Arabic *tafʿīla* foot—especially given that the Swahili metrical meta-vocabulary is almost entirely borrowed from Arabic. Yet, they do no such thing. When confronted with the physical reality of the manuscripts, they do not invent exogenous accentual rhythms or import Arabic quantitative rules that do not exist; instead, they faithfully and strictly describe the endogenous syllabic architecture of Swahili, a Bantu language. They remained loyal to the observable reality of the texts, explicitly and rigorously utilizing the exact endogenous terminology of the Swahili master poets themselves, documenting the architecture exclusively through the principles of *vina* and *mizani*.

This unwavering deference to the observable text and to Swahili vocabulary constitutes a major epistemological transgression within a disciplinary field heavily guarded by the accentualist trend. Scientific objectivity demands strict reflexivity in this context, particularly regarding the contradictory nature of the institutional pushback against empirical and data-driven metrics. A profound logical paradox emerges here: the very same Orientalist pioneers who centered their work on Bantu metrical architecture could simultaneously be disqualified as "Afrocentrist"—a logical impossibility, since an Orientalist bias and an Afrocentric bias are, by definition, mutually exclusive postures.

This contradiction exposes a flawed underlying syllogism, which can be stated formally as follows:

- **P1.** The accentual school constitutes the exogenous, imported paradigm against which any contradiction of it is measured.
- **P2.** An authentically Afrocentric posture, by definition, rejects exogenous frameworks imposed upon African material.
- **C (fallacious).** Therefore, any posture that contradicts the accentual school is an Afrocentric posture.

The conclusion does not follow, because P2 describes only a necessary surface resemblance, not a sufficient condition: the mere act of contradicting an exogenous model that seeks to erase or superimpose itself upon a demonstrably endogenous system does not, in itself, constitute an Afrocentric posture—it may just as well constitute a strictly empirical one. Disqualification as "Afrocentrist" therefore has little to do with actual ideological positioning, but is instead triggered the moment a researcher fails to pledge allegiance to the a priori paradigm of the accentual school. Consequently, this same reductive charge of Afrocentrism applies just as readily to any contemporary research that is not remotely Orientalist, immediately branding as ideologically suspect any endeavor that diverges from the accentualist line—including our own.

It is essential to further dispel this mirage of Afrocentrism that this logic creates. Our stance is neither a political manifesto nor a subjective ideological posture, and we harbor a profound, fundamental aversion to ideological dogmatism masquerading as scholarship. Rather, our approach is strictly pragmatic, Ockhamian, and governed by data-driven, formal and scientific methodologies. By evaluating the text through formal and experimental methods, this research aims to demonstrate that the syllabic structure is so empirically robust that it renders any exogenous projection entirely unnecessary.

By pragmatic, we mean that this methodology evaluates poetic theories exclusively by their practical application and their observable, predictable results in the material world of the texts. By Ockhamian, we refer to the absolute application of the principle of parsimony (Ockham's razor), which dictates that when confronted with competing hypotheses, the model requiring the fewest external, unproven assumptions—such as imported classical Greek trochees—is inherently and scientifically superior. Finally, our reliance on the experimental method mandates that any theoretical and formal framework must be rigorously tested against observable, quantifiable

data (Popper, 1959). If the empirical evidence contradicts the theory, the theory must be irrevocably discarded, regardless of its institutional prestige or the historical weight of its proponents.

A commitment to Popperian epistemology is incomplete, however, without a precise statement of the conditions under which the competing models would have to be abandoned. Scientific objectivity demands that no theoretical paradigm is immune to empirical refutation. It is precisely this explicit falsifiability that distinguishes the present approach from the accentual tradition it contests—a tradition which, tellingly, does not itself deny the syllabic substrate. To guarantee a perfectly symmetrical experimental protocol, we establish eight absolute conditions of falsifiability. For conceptual clarity at this stage of the demonstration, these parameters are presented through their structural definitions, prior to their full formal modeling (see Section 9.2.4 for details).

The first four conditions rigorously test the endogenous arithmetic and syllabic matrix, whose principles are formally modeled in Chapter 8 :

1. The absolute verification of the syllabic measures (line and hemistich syllable counts).

2. The spatial locking of local and global rhyme chains.

3. The syntactic impermeability of the metrical caesura.

4. The strict adherence to topological permutations and stanzaic chaining.

To ensure a truly symmetrical experimental protocol, Section 2.5 of this work formally models not one but two distinct exogenous accentual models: the historical accentualist paradigm inherited from Werner (1917-20) and most recently systematized by Vierke (2011) through her *gemischt-quantifizierend* (mixed-quantifying) system and cognitive *Grenzschema*, and the more recent diffusionist refinement proposed by Minerba (2025), who projects a "creolized trochaic" innovation onto the verse. Both models, notably, acknowledge the presence of the syllabic count in the corpus—Minerba himself states that isosyllabism characterizes every classical Swahili meter (Minerba, 2025, p. 156)—yet in both cases this concession is treated as a self-evident, almost trivial backdrop, something so obvious that it is assumed to require no further theorization, rather than as the generative principle of the verse itself. The syllabic layer is seen, but not investigated; and upon this

unexamined backdrop, an accentual rhythm is superimposed as though it constituted the true architecture at work. By subjecting these descriptive theories to mathematical reduction, we established the exact conditions for their falsification. The remaining four conditions apply the exact same empirical crucible to these exogenous accentual frameworks:

5. Testing the bipartite accentual axiom (the two-beat rule).

6. Testing the Grenzschema as an idealized cognitive prototype.

7. Testing the periodicity of asymmetrical stress distributions.

8. Testing the tetrasyllabic (four-syllable) foot against extreme metrical boundaries.

The falsification pursued in Chapter 9 therefore proceeds, symmetrically, against the formalized endogenous arithmetic and syllabic grid and against the accentual paradigm in both its historical formulation and its most recent and most sophisticated iteration. If the exogenous accentual paradigms fail to survive their own conditions of falsifiability during the testing of Chapter 9, they must be discarded as phonetic illusions and theoretical projections, or more simply errors. Conversely, should any of the four endogenous pillars fail empirically, this would not merely weaken the syllabic model: it would indicate that Swahili classical meter obeys other, as yet unidentified principles altogether.

This data-driven intellectual matrix does not originate in the subjective, often speculative realms of traditional literary hermeneutics. Instead, it is firmly anchored in the uncompromising demands of formal linguistics, formal logic, scientific methods and the principles of Open Science.

To empirically understand the architecture of a classical Swahili stanza following the endogenous metrical grid is to grasp a complex system governed by the syllabic metrics defined by the rhyme (*vina*) and the syllabic count (*mizani*). By applying the rigorous logic of regular expressions—a foundational concept in formal language theory used to define allowed sequences, spatial permutations, and strict boundaries within a string of data—we can perfectly map the exact constraints of this syllabic metrics that the Swahili masters actively engineered (see Chapter 8). Furthermore, the commitment to Open Science ensures that this formalization is not hidden behind esoteric academic jargon; the raw archival data, the digitized

manuscripts, and the resulting formal models are transparently provided, rendering the entire metrical hypothesis entirely falsifiable and accessible for independent verification. The primacy of data and material reality, together with formalization as a means of reducing ambiguity, constitutes the baseline epistemological requirement of any rigorous modern linguistic science.

Scientific objectivity demands that this exact same methodological rigor be applied symmetrically. We do not restrict this formalization to the endogenous system; rather, we subject the descriptive claims of the exogenous accentual models to an identical mathematical reduction (see Section 2.5). By translating their rhythmic hypotheses into strict, testable formal models, we ensure that both competing paradigms face the exact same uncompromising empirical crucible during the tests of Chapter 9.

This methodological foundation creates a profound, and perhaps irreducible, epistemological chasm between a strictly data-driven approach and the methodologies of scholars trained exclusively in classical literary theory. While a certain type of literary theorist might struggle to grasp the necessity of experimental verification—often prioritizing the conceptual elegance, or the historiographical lineage of a speculative accentual model (Vierke, 2011)—this framework is inexorably bound to the supremacy of empirical data. Material evidence, formal logic, and experimental validation will always supersede theoretical dogma.

The same epistemological posture this chapter defends for poetic meter finds an instructive parallel in a directly adjacent field of Swahili studies: population genetics. We invoke this case not as direct evidence for the syllabic hypothesis itself, but as a demonstration, within the same broad discipline, of how a diffusionist, ideologically-inflected narrative can be overturned by the sheer primacy of empirical data—precisely the intellectual posture this book asks its reader to adopt toward meter. For over a century, the Swahili coast was romantically and falsely cast, in a nineteenth- and twentieth-century diffusionist tradition, as a kind of transplanted, passively colonized "Al-Andalus"—a civilization understood as foreign to the African continent on which it stood. The ancient-DNA sequencing published by Brielle and colleagues in 2023 dismantled this narrative not through argument but through data. Their analysis of medieval and early-modern coastal towns found genome-wide ancestry to be substantially and genuinely admixed—on average close to an even split between sub-Saharan African and Asian ancestry, though the

exact proportion varies considerably from one individual to another—with this admixture beginning by roughly AD 1000, coinciding with the large-scale regional adoption of Islam (Brielle et al., 2023). However, this admixture proves to be sharply sex-biased: the African component is traced overwhelmingly to female ancestors, while the Asian component—drawn primarily from Persia, with a smaller Indian contribution—is traced overwhelmingly to male ancestors (Brielle et al., 2023). The picture, in other words, is not one of a foreign population trivially absorbed by a dominant African substrate at the level of aggregate ancestry; the Asian genetic contribution is far too substantial for that framing to hold. But it is precisely at the level of lineage, rather than of aggregate proportion, that the data speaks most clearly: Persian and Indian men married into African maternal lines, and their descendants were raised within a kinship structure that remained, at its reproductive and social foundation, African and matrilineal. The borrowed element—here, a paternal genetic and religious one; in the poem, a borrowed metrical meta-vocabulary—was absorbed into a substrate whose own generative and organizing logic never ceased to be African. Genetic ancestry, exactly like metrical terminology, can be substantially borrowed at the surface while the deep structure that organizes it, whether kinship or verse, remains endogenous.

This same distinction—between a borrowed surface vocabulary and an endogenous deep structure—is precisely what the formalization and experimental analysis of Chapters 8 and 9 will establish for Swahili verse. When confronted with the mechanics of Swahili verse, experimental observation will reveal, as Chapter 9 sets out in full, a strictly syllabic metrical system—based entirely on the absolute count of syllables (*mizani*) and terminal homophony (*vina*)—that is entirely devoid of any metrical stress pattern.

It is precisely upon this empirical foundation that Ockham's razor must be applied. Since the structural reality of the syllabic system will be shown to perfectly and sufficiently account for the poetic form, its rare failures being the defective lines (*guni*) that the tradition itself names, postulating an additional, imported accentual framework multiplies theoretical entities without necessity. Consequently, the principle of parsimony dictates that we must reject these superfluous exogenous variables and favor the only experimentally verified hypothesis: that of an endogenous, self-sustaining generative matrix based on syllabic metrics.

Acknowledging the historical reality of Indian Ocean trade, genetic *métissage*, and Islamic integration is a historiographical necessity, yet it by no means justifies the perpetuation of a nineteenth- and twentieth-century diffusionist model that romantically and falsely paints the Swahili coast as a passive, transplanted "Al-Andalus". The religious and mercantile influences from the Arabian Peninsula and the Persian Gulf are undeniable, but the operative, generative core of the poetic system remains local, syllabic, and autonomous.

Applying this strict experimental protocol and formal logic to the dominant accentual theories will yield unequivocal results: the accentual model imposed on Swahili poetry, in both its historical form (Werner, 1917-20; Vierke, 2011) and its most recent refinement (Minerba, 2025), will be shown to be formally and empirically unsound. The data, once formalized and tested, will demonstrate that the stress-based system is unsupported by the corpus and factually incompatible with the classical corpus, while the syllabic system alone accounts for its architecture, the residue being the defective lines (*guni*) that the tradition itself names.

However, in the institutional academic arena, the potential invalidation of a long-standing paradigm by empirical data does not automatically guarantee a serene scholarly debate. When a data-driven model—supported by endogenous terminology (see Chapter 4) and poised for empirical testing (see Chapter 9)—challenges a century of exogenous rhythmic projection, it inevitably confronts institutional inertia. In the history of science, the transition between paradigms is rarely without epistemological friction. While exposing the fragility of the accentual system through formal modeling ideally invites a rigorous and objective scholarly controversy—the very catalyst for a true paradigm shift—it can also trigger a structural reflex aimed at preserving an established conceptual framework. The introduction of falsifying empirical data tests the epistemological probity of a discipline, revealing whether it will adapt to new evidence or retreat into orthodoxy.

A Note on Institutional Reception

This potential friction highlights the complex sociology of academic networks. While such circles of influence strive to operate as open, progressive spaces for scientific falsifiability, they can also function as conservative ecosystems where the preservation of an inherited consensus inadvertently overshadows empirical anomalies. To challenge this established framework with verifiable data, formal models, and the undeniable voices of the Swahili practitioners themselves is to test the flexibility of these academic structures and their capacity to rethink how African texts are analyzed.

The "Procrustean Bed" is therefore best understood not merely as a passive error in metrical analysis, but as a structural mechanism that has historically maintained a specific academic order. While Swahili poets were historically marginalized through the application of an alien scansion by the accentualist school, the introduction of data-driven research today offers an opportunity for a necessary epistemological rupture. Whether this empirical approach is met with a constructive paradigm shift or with institutional resistance remains to be seen, but the scientific imperative requires that the data lead the way, unapologetically challenging any theoretical framework that may be facing its own obsolescence.

Chapter 8. The Sovereign Grid: The Endogenous Logic of Swahili Syllabic Meter

To model the classical Swahili metrical system accurately, we must set aside the accentual notations imported into its study, which describe a line as a sequence of stressed beats. The classical Swahili poet (*mtunzi*) works within a rigid matrix governed exclusively by the syllable count (*mizani*) and the rhyme (*vina*).

8.1. The Epistemological Framework: From Data to Formalization and Tests

The formalization presented in this chapter aims to extract underlying structural regularities from the classical tradition and translate them into a rigorous generative model. Fundamentally inductive and data-driven, this approach operates through abstraction: it synthesizes the structural principles explicitly articulated by the poets and scholars writing in Swahili with the empirical reality observed in the manuscripts. The primary objective is to bring a strict analytical order to experimental verification, thereby significantly increasing the falsifiability of our metrical hypotheses. By defining exact structural boundaries, the model provides precise formal matrices and corresponding algebraic equations through which any classical poem can be analyzed. However, within a data-driven epistemology, independent empirical data remains the absolute judge of the theoretical model. We do not evaluate the validity of the poem based on its adherence to our rules; rather, the formalized rules themselves are tested against new texts, in strict accordance with the specified falsification conditions.

Methodologically, this formalization is deeply inspired by formal language theory and theoretical computer science—specifically the logic of regular expressions. Originally conceptualized to map the precise syllabic constraints (*mizani*) and rhymes (*vina*) of Swahili verse, our logic provides a powerful syntax for defining allowed sequences, permutations, and boundaries within a metric string.

However, a metrical theory must remain legible and useful to a literary and philological audience. To bridge the gap between formal abstraction and literary analysis, this chapter adopts a dual presentation. First, we provide a spatial, two-dimensional representation of the matrices. These visual grids are specifically designed to map the stanzaic architecture as it appears in romanized printed texts, offering an intuitive typographical layout for the general scholarly public. Beneath these grids, we push the formalization further by extracting the corresponding algebraic equations. This symbolic reduction maximizes the model's concision and predictive power.

This algebraic approach is close to the formal metrics of the French linguist Benoît de Cornulier (1982, 1995). Cornulier's framework suits this endeavor because it defines meter not as a rhythmic flow but by the equality of syllable counts between lines, and by constraints on fixed positions such as the caesura—an apt lens for the isometry of classical Swahili poetry.

It is important to note that the formal matrices and algebraic equations modeling classical Swahili verse were initially developed independently, driven by both the endogenous descriptions of the tradition and early empirical observation, prior to our engagement with Cornulier's theoretical work. Moreover, to capture the complexity of Swahili poetics, this model goes beyond Cornulier's framework on one point: the formalization of rhyme chains (*vina*) according to their scope, distinguishing local stanzaic variables from global poem-length constants. By distinguishing between local stanzaic variables and global poem-length constants, this augmented model successfully renders the unique generative constraints of the Swahili tradition.

8.2. Formal Notation and Symbolic Alphabet of the Metrical Grid

To formalize the Swahili meter, the following symbols are defined; they state the axioms of the written grid in a notation close to that of regular expressions:

L: Represents a discrete formal line of verse *(mshororo)*.

L*n*: Represents the sequential rank or position of the line within a stanza (*ubeti*).

n_{max}**:** The parameter representing the upper bound of the possible values for L*n*; its value defines the meter's formal limit.

***H*:** The number of hemistiches (*vipande*) per line L.

[*N*]: The absolute number of syllables (*mizani*) within a specific segment or hemistich.

***S*:** The total syllable count (*mizani*) per line L. It is defined as the sum of all [*N*] segments within the line (*S* = [*N1*] + [*N2*] ... + [*NH*]).

> (hierarchical dominance between two forms of one variable): Placed between two indexed forms of the same metrical variable, this symbol records that the form on its left is the stronger condition and the one tested first. Written $\mathbf{F^s > F^v}$, it asserts two things at once. Logically, satisfying the left-hand form entails satisfying the right-hand one, while the converse does not hold: the two are nested and not alternative, and a corpus that meets the stronger condition automatically meets the weaker. Procedurally, the order of testing follows the symbol: the left-hand form is confronted with the corpus first, and the right-hand form is examined only where the left-hand one has been contradicted, the contradicting lines being reported. The symbol is therefore not decorative. It prevents a weaker form from being declared as a convenience, since declaring it presupposes a published failure of the stronger.

a, b, c... (lowercase) - Local rhymes (*vina*): Whose scope is limited to the individual stanza, and which change from one stanza to the next, binding the lines of a single strophic unit. A matrix is written with the variable unresolved, as *a*: the model states the parameter, and the corpus determines its form.

Inductive modeling of the classical corpus establishes that a local rhyme is attested in two forms, which stand in the hierarchy:

$a^s > a^v$

a^s, the *syllabic local rhyme*, requires that the lines it binds close on the same full consonant-vowel sequence, onset and vowel alike.

a^v, the *vocalic local rhyme*, requires the identity of the terminal vowel alone, the syllabic onset being free to vary: closings on *-a*, *-ya* and *-wa* satisfy it, since all three end on the same vowel.

The two are not alternative conventions on the same level. a^s is the rule of the tradition. a^v is a license, taken where the lexicon offers no syllabically rhyming word of the required meaning and measure, and it is by its rarity that it is recognized as one: a form employed in a handful of stanzas out of forty is a derogation from a rule, not a second rule.

Because a local rhyme is renewed at each stanza, the license may be taken in one stanza and not in the next, and the result is therefore reported as a count of stanzas. A corpus is described as employing a^s

in so many stanzas and a^{v} in so many, and the second figure measures the lexical pressure the composer met, stanza by stanza.

Each form is falsifiable on its own terms. A stanza declared a^{s} is contradicted by any bound line departing from the full syllable; one declared a^{v} is contradicted by any bound line departing from the vowel. Because the order of testing is fixed by the symbol, a stanza is never declared a^{v} without the failure of a^{s} having been established and the contradicting lines named.

X, Y, Z... (uppercase) - Global rhymes (*vina*): Acting as macro-structural constants, they anchor specific structural boundaries — the terminal rhyme *bahari*, for instance — across every stanza throughout the entire poem. A matrix is written with the constant unresolved, as X: the model states the parameter, and the corpus determines its form.

Inductive modeling of the classical corpus establishes that a global constant is attested in two forms, which stand in the hierarchy:

$X^{s} > X^{v}$

X^{s}, the *syllabic global rhyme*, requires the identity of the full consonant-vowel sequence at the anchored boundary: every anchored line of the poem closes on the same syllable, onset and vowel alike.

X^{v}, the *vocalic global rhyme*, requires the identity of the terminal vowel alone, the syllabic onset being free to vary.

The two are not alternative conventions on the same level. X^{s} is the rule of the tradition. X^{v} is a license, taken where the lexicon offers no syllabically rhyming word of the required meaning and measure, and it is by its rarity that it is recognized as one.

Because a global constant governs every anchored line of the poem at once, a single license anywhere resolves the constant for the whole work, and the result is therefore reported for the corpus as a whole rather than stanza by stanza.

Each form is falsifiable on its own terms. A corpus declared X^{s} is contradicted by any anchored line departing from the full syllable; one declared X^{v} is contradicted by any anchored line departing from the vowel. Because the order of testing is fixed by the symbol, a corpus is never declared X^{v} without the failure of X^{s} having been established and the contradicting lines named.

| (Vertical bar): The disjunction operator. It indicates a structural alternative or a mutually exclusive choice at a specific metric boundary.

An expression such as ([8,*a*]+[8,*b*]) | ([8,*b*]+[8,*X*]) signifies that the poetic grammar allows the position to be legally instantiated by either the line ([8,*a*]+[8,*b*]) OR the line ([8,*b*]+[8,*X*]).

+: The abstract formal marker for the metric caesura separating hemistiches (*vipande*) within a single line.

(*): The graphic separator utilized in Ajami manuscripts. The drawn markers are identified as *intahā* symbols[68], which are strategically placed to mark the caesuras (Parkar, 2021). It materializes the caesura between lines (*mishororo*) in *utenzi* and can also delineate the caesura between hemistiches (*vipande*) in other meters. It is essential to emphasize that scribal practice is not standardized; therefore, the notation is enclosed in parentheses to indicate that the physical marker is optional and exhibits significant graphical variation (e.g., spade-shaped symbols, three dots arranged in a triangle, stylized stars, flowers, etc.).

//: The abstract formal marker for the external metric boundary, indicating the end of a full line (*mshororo*) and the structural transition to the following line within the stanza.

· **(Central dot):** The formal concatenation operator. It denotes the strict sequential chaining of structural units to generate a higher-level form. Its specific scope is relative to the metrical model being formalized: within the syllabic model, it concatenates independent lines to construct the complete stanza.

[68] The term *intahā* (انتهى) is a Form VIII Arabic verb derived from the triconsonantal root *nūn-hā-yā* (ن ه ي), meaning "to come to an end, terminate, or reach a limit". In manuscript culture, it functions literally as an "end mark." As Parkar (2021) observes, graphical elements of this kind are not merely intended to beautify the text; they possess a definitive cognitive and structural function, designed to facilitate both the reading and the memorization of the metric architecture.

8.3. The Utenzi Meter

The *utenzi* (or *utendi* in northern dialects) is the engine of narrative and historical epics[69]. It is designed for maximum density and narrative stamina. Its foundational architecture is a stanza (*ubeti*) composed of four short lines (*mishororo*) devoid of any internal caesura.[70]

Before detailing the formal architecture of this meter, a brief methodological clarification is required regarding the epistemological flow of this chapter. The formal models presented in the following subsections (such as the matrices and algebraic equations) represent the final results of our empirical investigation of primary sources, not a priori theoretical assumptions. For pedagogical clarity, these abstract results are presented first, immediately followed by the empirical study of the graphic formalizations (the *ajami* manuscripts and romanized editions) that generated them.

The resulting abstract model is systematically presented through two complementary tools:

1. The Generative Matrix: A spatial representation of the stanza's architecture. It is important to acknowledge an inherent typographical bias here: for the sake of legibility within a modern printed book, these matrices are inevitably displayed using a vertical, romanized layout (one *mshororo* per line).

2. The Algebraic Equation: The ultimate symbolic reduction of the matrix. Unlike the matrix, this linear equation is entirely medium-agnostic. It transcends the romanized spatial bias and applies universally to both the vertical layout of Latin print and the continuous horizontal layout of endogenous *ajami* manuscripts.

This specific epistemological structure will govern the analysis of all three classical meters formalized in this chapter (*utenzi*, *wimbo*, and *shairi*). However, because *utenzi* is the first meter introduced, the

---

[69] The *utenzi* is defined exclusively by its formal metric architecture, independent of its thematic content. While it is the traditional meter for epics, equating *tenzi* solely with epics is inaccurate; the formal grid accommodates diverse genres, ranging from didactic morality like the *Utenzi wa Mwana Kupona* (Allen, 1971) to theological adaptation like *Utenzi wa Zaburi* "Poem of the Psalms" (Mnyampala, c.1965), among various other themes.

[70] While the late Sheikh Ahmed Nabhany also identifies a two-line layout (Style A) for the *utenzi* meter, this formal model deliberately adopts the four-line architecture (Style B) as the generative baseline. A full methodological justification for excluding the two-line layout from the primary formalization is provided in Section 8.3.2.

step-by-step extraction of the abstract models from the raw physical layouts will be demonstrated exhaustively in this section. For the subsequent meters, the exact same inductive logic applies, but the formal results will be presented more directly. The reader is therefore encouraged to closely observe the methodological progression in this section, as it serves as the foundational blueprint for understanding the remainder of the metrical formalization.

Focus 2: *Utenzi* and the Poetics of "Doing" — Translating the Word U*tenzi* in English

The Swahili noun *utenzi* derives from the verbal root *-tenda* meaning "to do" or "to act". Can we draw a conceptual parallel with the Greek *poiéō* or the Latin *gesta*? The analysis that follows establishes an etymological triangulation which resolves the conceptual scope of the term.

The noun *utenzi* belongs to the Swahili nominal class 11 (the *u-* class), which typically encompasses abstract concepts. In this case, the abstraction is built upon the Bantu verbal root *-tenda*. This semantic logic perfectly mirrors the conceptual foundation of the French term *poésie* (and its English counterpart poetry), which stems from the Ancient Greek verb ποιέω (*poiéō*), meaning "to make," "to manufacture," or "to create." Both the Swahili and the Greek terms thus point to the act of composition and artistic creation (*poïesis*). In the context of our current formal analysis, this specific semantic dimension perfectly echoes the structural sophistication and meticulous metrical crafting required to generate the *utenzi* matrices. It highlights the profound artistry of the poet, mastering the architecture of the verse through high intellectual creation.

However, the term *utenzi* also directly signifies "action" or "high deed." This invites a second, equally valid translation. Rather than viewing the work solely through the lens of Greek *poïesis* (the act of composing), *utenzi* echoes the Western medieval tradition of the *geste*. The term derives from the Latin *gesta*, the neuter plural of the past participle *gestus* (from the verb *gerere*: "to execute," "to accomplish"), literally designating "actions" or "things done" before it evolved to describe the epic narrative that recounts them. This semantic mapping is particularly fruitful for those classical *tenzi* (the plural form) dedicated to the narration of heroic exploits and historical battles—a major application for which the meter is perfectly suited.

The vitality of the *utenzi* rests precisely on this inherent versatility. It is not a fragmented concept, but a dual reality: it simultaneously

embodies the *poiéō* (the architectural crafting of the poem) and the *gesta* (the heroic or moral action being executed). Because it bridges both formal poetic sophistication and the narration of consequential actions, the meter cannot be confined strictly to the epic. This dual nature explains why it has historically provided a privileged space for moral, didactic, and religious reflection. From the bedrock of social normativity—of which *Utenzi wa Mwana Kupona* is the paradigm—to philosophical inquiries, historical chronicles, and even contemporary political jousts, this semantic versatility allows the meter to remain the sovereign receptacle of a Swahili discourse capable of embracing all registers of human experience.

It must be emphatically stated that this triangulation with Greek and Latin is employed strictly in a heuristic capacity. There is no genetic, historical, or etymological relationship between the Bantu root *-tend-* and these classical European terms. Rather, this comparative exercise serves solely as a conceptual bridge—a pedagogical tool to illuminate the dual nature of the Swahili *utenzi* through the familiar frameworks of Western literary tradition.

With this epistemological precaution established, this framework provides a highly precise guide for translation. The English rendering of *utenzi* must remain dynamically attuned to which of these two semantic poles a given text activates. When a work leans into the Greek *poiéō*—emphasizing the structural crafting, the art of composition, and the profound intellectual creation of the verse—the most accurate translation is simply 'poem'. Conversely, when the text operates within the realm of the Latin *gesta*—narrating heroic exploits, historical conquests, or consequential actions—the term 'epic' becomes fully justified. Acknowledging this duality frees Swahili poetry from taxonomic rigidities, ensuring that didactic, moral, or philosophical *tenzi* are correctly understood as masterfully crafted poems, rather than being erroneously shoehorned into an epic mold. A paradigmatic illustration of this is the canonical *Utenzi wa Mwana Kupona* (1858). In this renowned text, a mother imparts intimate moral and matrimonial advice to her daughter for a happy life. Translating this didactic work as 'The Epic of Mwana Kupona' is conceptually nonsensical; guided by the logic of *poiéō*, it must be rendered accurately simply as 'The Poem of Mwana Kupona'. [71]

[71] For transparency: the analysis above began as a synthesis that the present author wrote for the French Wikipedia entry on the *utenzi*, which it substantially expands and formalizes. See Utenzi. (2026, May 1).

8.3.1. Utenzi Meter: Generative Matrix for One Stanza (*ubeti*)

To isolate the generative engine of the *utenzi*, we first present its abstract formal model, standardly formatted here according to Western typographical conventions (vertical stacking, left-to-right).

Formal Parameters:
$n_{max} = 4$, $H = 1$, $N = S = 6$ or 8 or 11

The Matrix:
L1: [N, a] //
L2: [N, a] //
L3: [N, a] //
L4: [N, X] //

As this is the first time we translate a spatial matrix into a linear algebraic equation in this study, a brief methodological clarification is in order. This symbolic reduction condenses the stanza into a single generative string. The structural transition at the end of a line is marked by the double slash (//). Furthermore, to connect independent macro-structural blocks sequentially within the equation, we employ the central dot (·), which serves as the standard formal language operator for concatenation ("followed by").

Following our typological rules (where 'a' is the local variable and 'X' the global constant), the generative equation for the *utenzi* stanza can be elegantly condensed as follows:

$$\boldsymbol{Utenzi} = \mathbf{3}(\boldsymbol{N_a}//) \cdot \boldsymbol{N_X}//$$

(where N represents the chosen syllabic constant for the poem: 6, 8, or 11 syllables).

8.3.2. The Evolution of Graphic Reason: Two Spatial Instantiations

The equation that closes section 8.3.1 is an abstraction, which the page realizes in more than one form: historically in two, horizontal in the *ajami* manuscripts and vertical in romanized print, each altering the spatial apprehension of the poem. The matrices that follow render each of them as it appears on the page. They therefore set aside the abstract boundary markers of the symbolic alphabet, such as + and //,

and transcribe the demarcations between metrical units as the sources lay them out: line breaks follow the layout of the source, and adjacent segments are separated by a blank space or a typographical mark such as a comma. For *ajami* manuscripts, the generic character (*), defined in the symbolic alphabet, stands for the intahā ("end") markers, whose graphic form varies widely and which separate both hemistiches and lines. The aim is to show the stanza as the sources present it, without reducing it to algebraic notation.

1. The Ajami Manuscript Matrix

In the endogenous Arabic-script tradition, the *utenzi* stanza is frequently inscribed horizontally as a continuous ribbon across a single physical line of the page. The poet visually tracks the metric boundaries through graphic markers placed precisely after the rhymes. Graphically, the formal model is instantiated from *right to left* on a single continuous written line:

(*) [*X,* 8] :L4 (*) [*a,* 8] :L3 (*) [*a,* 8] :L2 (*) [*a,* 8] :L1

In this specific graphic layout, the generative algebraic equation of the stanza is mapped onto a strictly horizontal axis. Here, the four formal poetic lines (*mishororo*, represented by L_1 through L_4) are concatenated sequentially from right to left within a single physical graphic line (*mstari*). The abstract structural boundaries of the formal model are physically materialized on the page by the *intahā* markers (*). This demonstrates how a complete four-line metric architecture is visually compressed into a continuous, unbroken ribbon of text, entirely bypassing the need for modern typographical line breaks.

To empirically verify this continuous horizontal instantiation of the formal model in *ajami* script, the reader is invited to consult the digitized 1858 manuscript of *Utenzi wa Mwana Kupona* (MS 34882). As institutions like the School of Oriental and African Studies (SOAS) gradually open their historical archives to online public access, scholars can directly observe the unbroken right-to-left deployment of the four 8-syllable lines (*mishororo*) across a single physical graphic line (*mstari*). The manuscript is currently hosted by the SOAS Digital Collections and is available for online consultation.[72]

[72] Available on SOAS Digital Collections, *Utenzi wa Mwana Kupona* (item 1279415), last accessed August 27, 2026.

2. The Romanized Print Matrix

The late Sheikh Ahmed Nabhany describes two layouts for the romanized *utenzi* (Nabhany, 2006, p. [5]). It must be explicitly stated that the endogenous terminology used to describe these layouts is strictly Nabhany's own. In his theoretical framework, he defines Style A (*mtindo A*) as a layout structured around two poetic lines—which he specifically designates as *mishororo*—that map directly onto two physical graphic lines. Conversely, he defines Style B (*mtindo B*) as a layout structured around four poetic lines—which he also designates as *mishororo*—that map onto four individual graphic lines.

At this juncture, a conscious methodological decision must be stated. As demonstrated, our generative matrix formally models the *utenzi* stanza as a strict architecture of four *mishororo*. This invariant foundation successfully unifies the *ajami* layout (four *mishororo* concatenated across a single graphic line) with the romanized Style B (four *mishororo* distributed across four graphic lines).

To explicitly illustrate the spatial malleability and structural polymorphism of this model, we can map the base algebraic equation onto these two distinct graphic instantiations using a standard realization of eight syllables ($N = 8$):

The Symbolic Algebraic Equation

$$\mathbf{Utenzi} = \mathbf{3}(\boldsymbol{Na}\ //) \cdot \boldsymbol{NX}\ //$$

The *Ajami* Instantiation (Read Right-to-Left on a single line)

(*) [X, 8] :L4 (*) [a, 8] :L3 (*) [a, 8] :L2 (*) [a, 8] :L1

The Romanized Style B Instantiation (Read Left-to-Right, Top-to-Bottom)

L1: [8, a]
L2: [8, a]
L3: [8, a]
L4: [8, X]

Through these three equivalent representations, the polymorphism of the matrix becomes evident: a single abstract generative formula securely governs two radically different typographic realities without altering the underlying metrical architecture.

Consequently, the two-line layout of Style A is deliberately excluded from our primary formalization. This exclusion is an assumed operational choice, justified by two factors. First, the four-line layout

represents the dominant typographic standard in the vast majority of transcribed romanized texts. Second, and most importantly, the empirical testing for the *utenzi* meter conducted in Chapter 9 relies exclusively on the recorded reading of an L1 Swahili speaker (Fatuma Shaaban) from standard romanized materials. Because her test materials are formatted in this exact four-line disposition, aligning the formal model with the four-line *mshororo* structure is a methodological necessity.

2.a. Style A (*mtindo* A) — *Utenzi* with two lines *(mishororo)*

(Excluded from the formal model, but represented here for historical exhaustivity)

L1: [8, a] , [8, a]
L2: [8, a] , [8, X]

Shariff provides an example from a classical poem following Style A, albeit with a different syllable count (Shariff, 1988, p. 52):

L1: [6, a] , [6, a]
L2: [6, a] , [6, X]

2.b. Style B (*mtindo* B) — *Utenzi* with four lines *(mishororo)*

L1: [8, a]
L2: [8, a]
L3: [8, a]
L4: [8, X]

Abedi provides some examples following Style B, albeit with a different syllable count. He also acknowledges $S=8$. (Abedi, 1954, p. 30-32):

L1: [6, a]
L2: [6, a]
L3: [6, a]
L4: [6, X]

L1: [11, a]
L2: [11, a]
L3: [11, a]
L4: [11, X]

8.3.3. Methodological Note on the Data-Driven Model

It is essential to remember that this formal model of the *utenzi* is strictly data-driven, built upon an inductive methodology that requires a constant back-and-forth between abstract formalization and empirical reality. The generative matrix does not pre-exist the tradition; rather, it is directly derived both from the physical manuscripts and from the normative descriptions supplied by the poets themselves. Moving forward in this study, this abstract model will be consistently presented using the aforementioned romanized typographical conventions—specifically adhering to Style B—to maintain direct consistency with the physical format of the printed materials utilized in our empirical analysis. It is essential to remember that this formal model of the utenzi is strictly data-driven, built upon an inductive methodology that requires a constant back-and-forth between abstract formalization and empirical reality. The generative matrix does not pre-exist the tradition; rather, it is directly derived both from the physical manuscripts and from the normative descriptions supplied by the poets themselves. Moving forward in this study, this abstract model will be consistently presented using the aforementioned romanized typographical conventions—specifically adhering to Style B—to maintain direct consistency with the physical format of the printed materials utilized in our empirical analysis. Within this equation, N is a variable whose value is fixed for each poem. Abedi defines the *utenzi* as a line of fewer than twelve *mizani* (syllable measures) and documents three configurations, six, eight and eleven; these are the values retained here, and they generate the parameter $N \in \{6, 8, 11\}$. The parameter is therefore not read off the corpus tested in this study but taken from the endogenous definition, and the corpus is what tests it. Should attested poems fall outside this set, the matrix would have to be revised accordingly, as the inductive and iterative method of this study requires.

Moreover, this methodological principle extends beyond the utenzi and applies to all the meters formalized in this study. It is important to clarify that this formalization is never merely an ad hoc summary extracted from the specific datasets being tested. By actively integrating the general descriptions and structural rules articulated by the poets and scholars writing in Swahili, the matrices function as robust endogenous models. It is precisely this deductive final step—formalizing the rules stated by the tradition and subsequently applying them to the physical corpus—that gives the model its true evidential weight and protects the verification process from circularity.The parameter is therefore not read off our corpus but taken from the endogenous definition, and the corpus is what tests it.

Moreover, this methodological principle extends beyond the *utenzi* and applies to all the meters formalized in this study. It is important to clarify that this formalization is never merely an ad hoc summary extracted from the specific datasets being tested. By actively integrating the general descriptions and structural rules articulated by the poets and scholars writing in Swahili, the matrices function as robust endogenous models. It is precisely this deductive final step—formalizing the rules stated by the tradition and subsequently applying them to the physical corpus—that gives the model its true evidential weight and protects the verification process from circularity.

8.4. The *Wimbo* Meter

Although the noun *wimbo* literally means "song" (deriving from the verbal root *-imba*, "to sing"), this etymology should not imply that it is the only poetic genre intended for vocal performance. Just like the *utenzi* and *shairi* meters, the *wimbo* is inherently written to be sung. Because of the varied melodic executions of these sung performances, the *wimbo* is frequently and erroneously cited by accentual theorists as a domain of "free rhythm."

In reality, it is distinguished from other meters not by its musicality, but by the rigor of its metrical grid. By adhering strictly to the endogenous taxonomy provided by Ibrahim Noor Shariff in *Tungo Zetu* (1988, pp. 45–49), we uncover a highly complex, arithmetically unyielding syllabic architecture. The *wimbo* stanza is formally constrained by a strict limit of exactly three lines ($n_{max} = 3$); yet, within this inflexible boundary, the poet wields absolute architectural freedom.

8.4.1. The Standard Bicolumnar Matrix and Algebraic Equation (Symmetrical and Asymmetrical)

The most common *wimbo* utilizes two hemistiches ($H = 2$), but the poet (*mtunzi*) is free to define the syllabic weight of each side of the caesura. Shariff documents numerous variations (*mikondo*), including the canonical 8+8 (*nane kwa nane*), but also strictly asymmetrical grids such as 4+8 (*nne kwa nane*), 8+4 (*nane kwa nne*), 6+4 (*sita kwa nne*), and 8+5 (*nane kwa tano*).

Formal Parameters:
$n_{max} = 3$, $H = 2$, $S = N + M$ (allowing $N = M$ or $N \neq M$).

The Matrix:
$L_1: [N, a] + [M, b]$ / /
$L_2: [N, a] + [M, b]$ / /
$L_3: [N, a] + [M, b]$ / /

The General Algebraic Equation for the Bicolumnar *Wimbo*:

$$\boldsymbol{Wimbo_{bicolumnar} = 3(Na + Mb\ //)}$$

Note: Whether the measure is an even 6 + 6 or a highly uneven 8 + 5, the algebraic generator remains perfectly constant, locking the rhymes into an immutable spatial grid.

8.4.2. The Rupture Matrix: Vertical Asymmetry

The absolute sovereignty of the syllable count is demonstrated when the poet actively shifts the parameters of the grid *within* the same stanza. Shariff highlights a specific *mkondo* where the first two lines operate on a 2+8 or 4+8 grid, but the final line dramatically expands to a symmetrical 8+8 grid. A metric governed by a rhythmic "beat" would collapse under this sudden structural expansion; the Swahili syllabic grid simply absorbs it as a calculated algebraic shift.

Formal Parameters (e.g. the 2+8 or 4+8 shifting to 8+8 model):
$n_{max} = 3$, $H = 2$, $S = N + 8$ and N = 2 or 4

The Matrix:
L_1: [N, a] + [8, b] / /
L_2: [N, a] + [8, b] / /
L_3: [8, a] + [8, b] / /

The Algebraic Equation for the Rupture *Wimbo*:

$$\boldsymbol{Wimbo_{rupture} = 2(Na + 8b\ //) \cdot (8a + 8b\ //)}$$

8.4.3. The Monocolumnar Matrix (*Wimbo wa Zina Zitoto*)

Shariff documents a rare but crucial variation composed of a single, undivided 8-syllable block per line ($H = 1, S = 8$). In this *mkondo*, the internal caesura vanishes entirely, and the rhyme occurs exclusively at the terminal boundary. Shariff notes that the final line of this specific variation is often permuted to become the first line of the subsequent stanza.

Formal Parameters:
$n_{max} = 3$, $H = 1$, $S = 8$

The Matrix:
L_1: [8, a] / /
L_2: [8, a] / /
L_3: [8, a] / /

The Algebraic Equation for *Zina Zitoto*:

$$\boldsymbol{Wimbo_{monocolumnar} = 3(8a\ //)}$$

8.4.4. The Hyper-Fractionated Matrix

The architectural complexity of the matrix is most evident in the hyper-fractionated *wimbo*. Shariff documents grids broken into three hemistiches (e.g., 6+6+4) and even four hemistiches per line. For instance, he provides a remarkable example of a four-hemistich *wimbo* generated by a highly asymmetrical 6+3+4+2 grid (*sita kwa tatu kwa nne kwa mbili*). In such structures, the internal boundaries

of the matrix are dictated exclusively by the interlocking rhyme chains, creating a single poetic line that dynamically shifts from a six-syllable block to a three-syllable block, then to a four-syllable block, and finally to a two-syllable unit.

Formal Parameters for the 4-Hemistich *Wimbo*:
$n_{max} = 3, H = 4, S = 6 + 3 + 4 + 2$

The Matrix:
$L_1: [6, a] + [3, b] + [4, c] + [2, d]\ //$
$L_2: [6, a] + [3, b] + [4, c] + [2, d]\ //$
$L_3: [6, a] + [3, b] + [4, c] + [2, d]\ //$

The Algebraic Equation for the Hyper-Fractionated *Wimbo*:

$$\boldsymbol{Wimbo_{fractionated} = 3(6a + 3b + 4c + 2d\ //)}$$

The *wimbo* grid explicitly demonstrates the division of labor inherent in classical Swahili oral arts. The composer (*mtunzi*) resolves the formal equation on the page—whether calculating a simple 3(8a //) or a highly complex fractionated matrix. Once this formal blueprint is secured, the performer (*mghani*) is granted total melodic freedom (*mahadhi*). The singer may stretch vowels, introduce syncopation, or alter the tempo, precisely because the invariant *vina na mizani* grid guarantees that the verse will not collapse formally.

8.5. The Shairi Meter: From the 16-Syllable Standard to the General Formal Grid

The *shairi* meter is the architectural pinnacle of classical Swahili poetics, requiring absolute mastery over a complex internal balance. Due to a distinct corpus bias—driven by the wide dissemination of printed literature in the twentieth century—the *shairi* meter is frequently reduced to its most popularized prototype: the tarbia[73] (four lines, i.e., $n_{max} = 4$). While this specific form does not encompass the full structural diversity of the meter, it serves as the dominant

[73] The term *tarbia* has an Arabic etymology (*tarbīʿ*, تربيع), literally meaning "to make four" or "quadrature," designating a four-line stanza (King'ei & Kemoli, 2001, p. 16). It perfectly echoes the parameter $n_{max} = 4$.

matrix in published texts. Within this prototype, every single line *(mshororo)* is strictly composed of 16 syllables, obligatorily divided by a rhyme-generated caesura into two perfectly symmetrical 8-syllable hemistiches (*vipande).*

8.5.1. 16-Syllable Shairi Meter: Generative Matrix and Algebraic Equation for One Stanza (*ubeti*)

Formal Parameters:
$n_{max} = 4$, $H = 2$, $S = 16$ (where $S = [\ 8\ ] + [\ 8\ ]$)

The Matrix:
L1: [8, *a*] + [8, *b*] //
L2: [8, *a*] + [8, *b*] //
L3: [8, *a*] + [8, *b*] //
L4: ([8, *a*] + [8, *b*]) |
([8, *b*] + [8, *X*]) |
([8, *b*] + [8, *a*]) |
([8, *X*] + [8, *Y*]) //

The Algebraic Equation for the 16-syllable *shairi* stanza:

$$\boldsymbol{Shairi}_{\boldsymbol{16-syllable}} = \mathbf{3}(\mathbf{8}_{\boldsymbol{a}} + \mathbf{8}_{\boldsymbol{b}}//) \cdot \left((\mathbf{8}_{\boldsymbol{a}} + \mathbf{8}_{\boldsymbol{b}}) \mid (\mathbf{8}_{\boldsymbol{b}} + \mathbf{8}_{\boldsymbol{X}}) \mid (\mathbf{8}_{\boldsymbol{b}} + \mathbf{8}_{\boldsymbol{a}}) \mid (\mathbf{8}_{\boldsymbol{X}} + \mathbf{8}_{\boldsymbol{Y}})//\right)$$

Note: The structural alternatives in the final line (L_4) correspond to specific rhyme patterns, each explicitly designated by an endogenous Swahili term. The option ([8,*a*]+[8,*b*]) represents *kushuka* (a descending continuation of the stanza's main rhymes); ([8,*b*]+[8,*X*]) corresponds to *kupindua kina kimoja* (the permutation of a single rhyme); ([8,*b*]+[8,*a*]) illustrates *kupindua* (the complete inversion of both rhymes); and finally, ([8,*X*]+[8,*Y*]) designates *kurejelea* (the recurrence, where the final line repeats as a refrain throughout the poem).

8.5.2. Shairi Meter: General Generative Matrix for One Stanza (*ubeti*)

Drawing on Shariff's (1988, pp. 49–51) work, we now propose a more general matrix for the *shairi* meter, one that can model forms with asymmetrical hemistiches, stanzas exceeding four lines *(mishororo)*, and other structural variations.

Formal Parameters:
$n_{max} \geq 4$, $H = 2$, $S = N+M$ (where S = [N] + [M]) (allowing either $N=M$ or $N \neq M$)

The Matrix:
L1: [N, a] + [M, b] //
L2: [N, a] + [M, b] //
L3: [N, a] + [M, b] //
[...]
Ln_{max}: ([N, a] + [M, b]) | ([N, b] + [M, X]) | ([N, b] + [M, a]) | ([N, X] + [M, Y]) //

The General Algebraic Equation for the *shairi* stanza:

$$\boldsymbol{Shairi_{general}} = (\boldsymbol{n_{max}} - \boldsymbol{1})(\boldsymbol{N_a} + \boldsymbol{M_b}//) \cdot \big((\boldsymbol{N_a} + \boldsymbol{M_b}) \mid (\boldsymbol{N_b} + \boldsymbol{M_X}) \mid (\boldsymbol{N_b} + \boldsymbol{M_a}) \mid (\boldsymbol{N_X} + \boldsymbol{M_Y})//\big)$$

Note: The upper bound of n_{max} is left formally open ($n_{max} \geq 4$) to accurately reflect the structural flexibility explicitly codified in Swahili metrical theory. As Ibrahim Noor Shariff notes, while *shairi* stanzas of four lines are among the most well-known and frequently composed (*zenye kujulikana sana na kutungiwa sana*), a stanza consists of four lines "or more" (*mine au zaidi ya mine*) (Shariff, 1988, p. 49). He further specifies that the composer retains the sovereign freedom to expand the stanza incrementally to "five or six and so on" (*mitano au sita na kadhalika*) (Shariff, 1988, p. 50). This open parameter therefore records a generative capacity governed by the rules of the craft itself, rather than a theoretical license taken by the model. The formulation reflects a practical, incremental architectural expansion rather than a theoretical infinity. Like the syllabic variables N and M, the precise value of n_{max} is deterministically resolved by the empirical text in each case study.

8.5.3. The Evolution of Graphic Reason: Two Spatial Instantiations

1. The Ajami Manuscript Matrix

In this section, we focus exclusively on the 16-syllable *shairi* with octosyllabic hemistiches. For this example, the rhyme scheme is also fixed, as the primary objective is to demonstrate the *ajami* spatial layout.

To verify the structural correspondence between the abstract formal model and its primary instantiations, the reader is invited to examine the poem *Kiswahili* by Mahmoud Ahmad Abdulkadir. By comparing the *ajami* manuscript (Abdulkadir & Frankl, 2013, pp. 11–13) with the romanized text provided in the same article (Abdulkadir & Frankl, 2013, pp. 5–8)—both of which display a clear bicolumnar spatialization of the 16-syllable line—one can observe two distinct spatial transpositions of the same underlying formal logic. In both scripts, the structural integrity of the caesura is maintained through a deliberate central blank space, demonstrating that a spatial reality echoes the metric division. The 16-syllable line unfolds horizontally. The caesura [+], in the specific composition under examination, is a physical spatial reality on the page—a deliberate blank space that separates the right and left hemistiches into a bicolumnar layout, without the need for a drawn graphic marker. The poet (*mtunzi*) mentally and visually weighs the balance of the line across this central void.

Graphically, the formal model is instantiated from right to left, with defined parameters ($n_{max} = 4$, $H = 2$, $S = 16$, $N{=}M{=}8$):

[b, 8]	[a, 8] :L1
[b, 8]	[a, 8] :L2
[b, 8]	[a, 8] :L3
[X, 8]	[b, 8] :L4

Alternatively, the caesura [+] can be graphically rendered on the page. As Annachiara Raia (2020, p. 122) observes in some short devotional *shairi* poems she analyzes, each eight-syllable hemistich (*kipande*) is separated by three dots arranged in a triangle, marking the caesura between the hemistiches (*vipande*). Consequently, to accurately describe this specific *ajami* script, one should use our

formal symbol (*) between the hemistiches instead of a simple blank space:

[b, 8] (*) [a, 8] :L1
[b, 8] (*) [a, 8] :L2
[b, 8] (*) [a, 8] :L3
[X, 8] (*) [b, 8] :L4

However, the abstract model using the [+] symbol remains valid to represent the caesura, regardless of its graphic form.

2. The Romanized Print Matrix

2.a. The Canonical 16-Syllable Shairi

Defined Formal Parameters:

$$n_{max} = 4,\ H = 2,\ S = 16,\ N{=}M{=}8$$

Note : Four lines, symmetrical octosyllabic hemistiches

Romanized typography translates this foundational equilibrium by structuring the hemistiches into a bicolumnar layout, or by replacing the spatial caesura with a comma, or, worse, by breaking the 16-syllable line into two separate, vertically stacked 8-syllable lines mimicking the romanized *utenzi* layout[74].

This vertical stacking visually dismantles the original architectural intent of the classical poet. To preserve the structural integrity of the composition, one poetic verse (*mshororo*) must correspond exactly to one continuous graphic line (*mstari*). Consequently, the visual presentation of our abstract formal model deliberately adopts an unbroken, left-to-right romanized layout. The primary justification for this choice, however, is methodological: this unbroken format faithfully reflects the dominant editorial convention of modern romanized editions. Since the subsequent instantiations and empirical tests are strictly conducted on these printed versions, standardizing our visual presentation to match this established corpus ensures a consistent analytical framework.

[74] This is one of the structural contexts to which Shariff's (1988, p. 42) remark applies, noting that a verse line (*mshororo*) does not necessarily correspond to a physical line on paper (*mstari*).

When instantiated through this chosen unbroken layout, the four-line model reveals a clear internal dynamic. While the first three lines remain metrically homogeneous, structural variations emerge specifically in the fourth and final line, dictated by the different trajectories of the interlocking rhyme chains. To map this architectural flexibility, the following outlines the various topologies of these rhyming patterns, illustrating the primary spatial layouts found in published works.

2.A.A. KUPINDUA KINA KIMOJA — PERMUTATION OF A SINGLE RHYME (B ONLY)

The two layouts differ solely in the graphic materialization of the caesura, which is either explicitly marked by a comma or rendered spatially through a central blank space (Abdulkadir & Frankl, 2013, pp. 5-8; Shariff, 1988, p. 49).

L1: [8, a] , [8, b]
L2: [8, a] , [8, b]
L3: [8, a] , [8, b]
L4: [8, b] , [8, X]

Or

L1: [8, a]	[8, b]
L2: [8, a]	[8, b]
L3: [8, a]	[8, b]
L4: [8, b]	[8, X]

2.A.B. KUPINDUA — PERMUTATION (RHYME A AND B)

Nabhany (2006, p. [4]) and Shariff (1988, p. 49) describe the possible permutation (*kupindulia*) of rhymes *a* and *b* in line 4, without specifying whether this operates as a global or a local rhyme scheme:

L1: [8, a]	[8, b]
L2: [8, a]	[8, b]
L3: [8, a]	[8, b]
L4: [8, b]	[8, a]

2.A.C. KUSHUKA — DESCENDING: THE RHYME CHAINS A AND B CONTINUE WITHOUT PERMUTATION IN THE FINAL LINE

Shariff (1988, p. 49)

L1: [8, a] [8, b]
L2: [8, a] [8, b]
L3: [8, a] [8, b]
L4: [8, a] [8, b]

2.A.D. KUREJELEA — RECURRENCE: THE FINAL LINE REPEATS THROUGHOUT THE POEM

Shariff (1988, p. 50)

L1: [8, a] [8, b]
L2: [8, a] [8, b]
L3: [8, a] [8, b]
L4: [8, X] [8, Y]

A hapax can be found in Shaaban Robert's *Sanaa ya Ushairi* (Robert, 1972, pp. 43-44), in the poem *Bakia nyimbo*:

L1: [8, a] , [8, Y]
L2: [8, a] , [8, Y]
L3: [8, a] , [8, Y]
L4: [8, X] , [8, Y]

A variant of this scheme appears in *Swali langu ni ujane* (Robert, 1972, pp. 44-47):

L1: [8, a] , [8, Y]
L2: [8, a] , [8, Y]
L3: [8, a] , [8, Y]
L4: [8, X] , [8, Z]

Our abstract model integrates all these configurations, with the exception of Shaaban Robert's two latter schemes, which remain exceedingly rare[75].

[75] To be perfectly precise, there is a further theoretical justification for excluding Shaaban Robert's two aforementioned rhyme schemes from our abstract model: Robert himself never explicitly identifies them as *shairi*. In his treatise, he formally defines only the *"tenzi"* meter, which he characterizes primarily by its sustained rhyme repetition—essentially, its reliance on global rhyme constants. Because he omits a dedicated structural definition for the *shairi* meter, a taxonomic ambiguity arises. Since these two specific schemes rely on global constants, they might actually align

2.b. Other traditional sub-meters *(mikondo)*

The following section is deeply indebted to Shariff's seminal *Tungo Zetu*; even more than the preceding section, it is almost a translation, though here recast through the lens of formal modeling.

Here Shariff (1988, pp. 50–51) makes the variation of the parameters explicit: four-line stanzas may be built from symmetrical hexasyllabic hemistiches, or from asymmetrical patterns such as 4+8 or 6+8 syllables. He further notes that the number of lines is not fixed at four; it may be extended to five, six, or more, depending on the poet's will and skill. This is precisely where his critique of Kaluta Amri Abedi's *Sheria za Kutunga Mashairi* becomes visible: the standard 16-syllable models are not immutable laws, but flexible structures whose parameters may vary. Shariff writes:

> Up to this point, we have given examples of poems with stanzas of only four lines; however, the poet, if he so wishes, may increase the number of lines in each stanza, making them five or six, and so on. I hope that these few examples of song and poem meters are enough to awaken the student to the fact that there are many patterns of syllabic measure *[mizani]* within these meters. (Shariff, 1988, p. 51)[76]

At no point is there any reference to a beat or an accentual foot; instead, it becomes clear that we are dealing with a syllabic matrix of considerable expressive force. These variations justify our use of independent variables N and M to represent the number of syllables per hemistich, and of a maximum line count, $n_{max} \geq 4$, in the general abstract model.

---

more closely with Robert's own implicit conception of *tenzi*. Given this categorical uncertainty, they cannot be reliably integrated into a strict formal model of the shairi (Robert, 1972, p. xiii).

[76] Swahili original : "Mpaka hapa tumetoa mifano ya mashairi ya beti za mishororo minne tu, lakini mtungaji, akipenda, anaweza kuzidisha mishororo katika kila ubeti, ikawa mitano au sita na kadhalika. Natumai mifano hii michache ya bahari za nyimbo na mashairi inatosha kumzindua mwanafunzi kuwa kuna mipango ya mizani ya aina nyingi katika bahari hizo."

8.5.4. Transcending the *Shairi* Matrix: The 19th-Century Origins of *Mashairi ya Vidato*

This specific sub-meter, the *Mashairi ya Vidato*, stands as a definitive demonstration of the generative expansion of the Swahili syllabic grid.

During our initial doctoral research—which involved the complete digitization of Mathias E. Mnyampala's personal archives in Dodoma, in collaboration with his son Charles M. Mnyampala—the exact historical genesis of this form was still under investigation. To provide precise philological clarification on this matter, a specific extract of these unpublished archives, dedicated solely to the *Mashairi ya Vidato*, is now publicly referenced[77]. In our dissertation, relying on the prevailing biographical narratives and available family archives at the time, we explored the *Mashairi ya Vidato* form in depth (Roy, 2013b, pp. 370–379, including a structural formalization diagram on page 376). We noted then: "in the case of the *vidato* style, Mnyampala will be one of the initiators if not the main initiator, we do not know" (Roy, 2013b, p. 373). At the time, this assessment was shaped by the unpublished nature of the *Mashairi ya Vidato* folder and the assertion by Mnyampala's heir that the form was his father's innovation. Compounded by what was, in retrospect, a still-developing understanding of Swahili literary history, we had naturally concluded that the style was a 20th-century invention. This hypothesis, however, would ultimately prove to be mistaken.

According to his son, Charles M. Mnyampala, the poet Mathias E. Mnyampala was the primary initiator of this intricate sub-meter, a historical detail documented in his foreword of the posthumous publication *Maisha ni kugharimia* (Mnyampala, 2013a, pp. xi-xxvi). Furthermore, in Mnyampala's own unpublished archives, a marginal note on the poem *Jarida la Kiswahili*[78] explicitly defines the *Mashairi ya Vidato* as a "*tungo mpya*". Due to the semantic ambiguity of this phrase—which can mean either a "new metrical form" or simply a "new composition"—it was logically presumed that Mnyampala was its creator.

---

[77] To allow for independent verification of the primary manuscript sources debated here, a precise extract from the Mnyampala archives—strictly limited to the relevant *Mashairi ya Vidato* folios (myv5 and myv12) and their original folder cover (myv0a)—has been made publicly available as digital facsimiles:

Mnyampala, M. E., Roy, M.& Mnyampala, C. M. (1966). Primary Archival Extracts: Mathias E. Mnyampala's Unpublished "Mashairi ya Vidato" Manuscript (1966-1967). Zenodo.

Available at: https://doi.org/10.5281/zenodo.21362077 (Accessed: August 28, 2026).

[78] See archival code myv5.jpg in the previously cited digital archive.

However, the structural architecture of this meter actually dates back to the 19th century. We owe a debt of gratitude to the research of Anna N. Kyamba (2022) and Professor Madumulla, whose insights prompted a re-examination of the classical sources. As Kyamba points out, following a personal communication with Madumulla, the *Mashairi ya Vidato* form significantly predates Mnyampala's era. Following the integration of Kyamba's information, our own comprehensive verification of the *Diwani ya Muyaka bin Haji al-Ghassaniy* (Hichens, 1940) confirms that this metric form was actively utilized in the 19th century. The *Mashairi ya Vidato* structure is perfectly executed in the poem *Uwongo* "lie", authored by the classical poet Muyaka bin Haji and located on page 92 of his *Diwani*.

The analysis presented by Kyamba provides a valuable corrective by establishing the 19th-century precedence of the *Mashairi ya Vidato* form. Yet, paradoxically for a specialist in intertextual theory, her study contains significant methodological and philological anomalies that require clarification. While she accurately cites the poem *Uwongo* (which she transcribes as *Umwongo*) on page 92 of the *Diwani ya Muyaka* (Hichens, 1940, pp. 92-93), she erroneously claims that another *vidato* poem titled *Uaminifu* ("honesty") appears on page 1 of the same *Diwani*. In reality, the page 1 of *Diwani ya Muyaka* features the 20th-century prefatory homage by Mwalimu Sikujua bin Abdallah al-Bataway. *Uaminifu* is, in fact, an unpublished *Mashairi ya Vidato* poem authored by Mathias E. Mnyampala himself, explicitly documented in his primary archival extracts and dated July 18, 1966, on the original manuscript sheet.[79] By confusing Mnyampala's unpublished work with the 20th-century introductory pages of Muyaka's *Diwani*, she misattributes the poem's authorship[80].

---

[79] See archival code myv12.jpg in the previously cited digital archive.

[80] Additionally, there are notable bibliographic omissions in her article. Her analysis structurally mirrors the exact progression—*Mtiririko*, *Msisitizo*, *Majibizano/Ngonjera*, and *Vidato*—laid out by Charles M. Mnyampala in the preface to the 2014 edition of *Maisha ni Kugharimia*. She references "Roy (2014)" and "Charles Mnyampala (2014)" within her text to frame Mnyampala's innovations, yet the corresponding entry (Mnyampala, 2014) is unexpectedly absent from her bibliography. It is important to clarify that the 2014 text she relies upon is a standard global edition—published without the color photographs included in the foundational 2013 scholarly edition—that contains the exact same text (Mnyampala, 2013a). While Kyamba naturally cannot be faulted for omitting the open-access counterpart (Mnyampala, 2013b)—since this version was only recently shared on Zenodo in the spirit of Open Science after this Tanzanian literary treasure formally entered the public domain—the lack of a full bibliographic entry for the 2014 edition she explicitly uses creates a regrettable gap in scholarly traceability. Rather than deliberate academic

To fully understand the manuscript's timeline and properly attribute the paternity of this form, one must examine the philological history of the text. Muyaka's poems, which had long been transmitted orally, were collected and written down in the first half of the 20th century by a working group that included the poet and teacher Mwalimu Sikujua bin Abdallah al-Bataway from Mombasa, along with Sheikh Mbarak Ali Hinawy and William Hichens (Noronha, 2009). These poems were then edited and published by Hichens. Mwalimu Sikujua himself authored the introductory praises featured on the very first page of the manuscript. As a 20th-century scholar and poet, Sikujua deliberately composed this prefatory untitled poem[81] using the exact *Mashairi ya Vidato* metrical grid.

From a literary criticism standpoint, this is a profound piece of evidence: Sikujua utilized this highly complex, architecturally demanding structure to render a stylistic homage to Muyaka. It highlights that the *Vidato* sub-meter was recognized by 20th-century Swahili intellectuals as an elevated, masterful form, ideal for a tribute to a great master. Therefore, historical and philological evidence designates Muyaka bin Haji as the true father of the *Mashairi ya Vidato*.

Ultimately, while the generative genesis of this complex metrical grid firmly belongs to the 19th-century Mombasa poet Muyaka bin Haji, Mathias E. Mnyampala brilliantly mastered its execution. Metrically, Mnyampala's unpublished poem *Uaminifu* ("honesty") replicates the precise *Mashairi ya Vidato* structure of Muyaka's previously cited *Uwongo* ("lie"). Beyond this formal continuity, the two compositions are also bound by a striking thematic antithesis, establishing a clear semantic opposition between moral integrity and deceit. Across an interval of more than a century, Mnyampala appears to be directly responding to Muyaka.[82].

---

obfuscation, this omission appears to be an unfortunate editorial oversight, perhaps reflecting the publication of an unfinalized draft. Consequently, both her direct 2014 source and the foundational 2013a editions remain uncited.

[81] For the reader's clarity at this point in the demonstration, it should be noted that this prefatory poem, composed by Mwalimu Sikujua using the *Mashairi ya Vidato* metrical grid, is untitled (Hichens, 1940, p. 1) and is an entirely distinct composition from Mnyampala's *Uaminifu*.

[82] To avoid any ambiguity, as detailed in the main text: while Anna N. Kyamba correctly identified—with Professor Madumulla—Muyaka's 19th-century authorship of this metrical structure, the attribution of *Uaminifu* itself to Muyaka remains an error. Muyaka authored only *Uwongo* (Hichens, 1940, pp. 92-93).

Having established the historical and thematic dialogue between these poets, we can now turn to the structural formalization of this intricate sub-meter. It is important to emphasize that this generative expansion does not invalidate or modify the general classical matrix established above. Rather, the *Mashairi ya Vidato* operates as a highly complex internal expansion of the standard *shairi* structure, remaining strictly faithful to the endogenous *vina na mizani* framework. If we were to isolate a single *Mashairi ya Vidato* stanza and strip away its macro-structural context, it would perfectly mirror the classical *Kupindua kina kimoja* model described above, operating securely within standard parameters.

However, the innovation lies in "transcending" the traditional isolation of the stanza by introducing a structural interlocking. In this structure, the rhyme of the second hemistich of lines 1, 2, and 3 in stanza U_i permutes to the first hemistich position in the fourth line of the same stanza. This specific rhyme then functions as the first hemistich rhyme for the lines 1, 2, and 3 of the subsequent stanza U_{i+1}. The stanzas are thus structurally chained together in sequential pairs across the entire poem through a cascading permutation of the rhyme chains.

1. Shairi Meter: The Stanzaic Chaining Matrix and Its Algebraic Equation

To accurately model this stanzaic chaining, we must abandon the static variables $(a,\ b)$ used for classical, independent stanzas and introduce an indexed mathematical sequence of variables $R = (r_0, r_1, r_2, \dots, r_k)$. For any stanza of index $i \geq 1$, the stanza inherits the first-hemistich rhyme r_{i-1} from the rhyme migration of the preceding stanza (or from an initialization variable r_0 for the initial stanza) and introduces a new second-hemistich rhyme r_i.

Formal Parameters:
nmax = 4, H = 2, S = 16 (where S = [8] + [8])

The Matrix for Stanza U_i:
L_1: $[8, r_{i-1}] + [8, r_i]$ / /
L_2: $[8, r_{i-1}] + [8, r_i]$ / /
L_3: $[8, r_{i-1}] + [8, r_i]$ / /
L_4: $[8, r_i] + [8, X]$ / /

The Algebraic Equation for Stanza U_i:

$$\boldsymbol{U_i = 3(8r_{i-1} + 8r_i\ //) \cdot (8r_i + 8X\ //)}$$

To model the complete poem *P* of *k* stanzas, the equation defines a sequential concatenation over the index *i*. This formalization shows that the classical Swahili grid can function as a highly sophisticated generative algorithm capable of high structural complexity.

$$\boldsymbol{P = \sum_{i=1}^{k} [3(8r_{i-1} + 8r_i\ //) \cdot (8r_i + 8X\ //)]}$$

2. On the Etymology and Structural Logic of the Mashairi ya Vidato

The nomenclature of the *Mashairi ya Vidato* requires careful philological distinction. While the term *kidato* is often associated in modern Swahili with the secondary meaning of a "class in secondary school," such an interpretation is extraneous to the meter's internal logic. The primary etymological root of *kidato* refers to the notches or rungs carved into the trunks of palm trees to facilitate scaling—a literal "step" or "ladder." This etymology is indispensable for understanding the generative architecture of the meter: each stanza functions not as an isolated unit, but as a structural "rung" or "notch" within an ascending assembly. By adopting this term, Mnyampala explicitly announces the ladder-like architecture of the poem, where the *structural interlocking* of stanzas allows the text to "scale" vertically, step by step, through the sequential migration of the rhyme across cascading pairs of stanzas.

3. A Classical Blueprint: Illustrating the Vidato Mechanism through Muyaka's Uwongo

While the structural matrix, its corresponding algebraic equation, the subsequent summation algorithm, and the etymological framework detailed above are indispensable for defining the generative logic of the meter, such strict formalization can naturally be quite demanding for the reader. To gracefully transition from theoretical complexity to tangible literary beauty, it is essential to observe this intricate architecture in practice.

Below is the complete text of *Uwongo* ("lie"), composed by the 19th-century master Muyaka bin Haji al-Ghassaniy (Hichens, 1940, pp. 92-93). As the earliest known application of the *Mashairi ya Vidato*, it

materializes the "ladder" mechanism. The reader can visually trace how the stanzas are structurally chained together in sequential pairs through the migration of the rhyme.

Taking the first stanza as a starting point, observe the final rhyme of the first three lines (hemistich 2, ending in *-ye*). In the fourth line, this rhyme elegantly permutes to the internal position (hemistich 1). Then, the interlocking cascade occurs: this exact same *-ye* rhyme propagates into the following stanza, becoming the internal rhyme (hemistich 1) for its first three lines. This dynamic cycle repeats throughout the poem: the new final rhyme of the second stanza (ending in *-ne*) permutes to its fourth line and cascades into the third stanza, and so forth, step by step down the structural ladder.

Furthermore, this composition is not merely an illustrative historical example. It constitutes a foundational piece of our empirical corpus and will be tested against its formal model in Chapter 9.

Uwongo
Karatasi nakutuma, leo unitumikiye,
Kabalegheshe kalima, uzidi kusema naye,
Kudanganyana si kwema, kulla jambo kwa ndiaye,
Yalekeze yalekeye, tus'ambiane uwongo.

Yalekeze yalekeye, waolezi wayaone,
Tujivune khatimaye, mitima yetu isone,
Hakuna liso nduuye, mambo pia yafenene,
Sitaki tudanganyane, na kwambiana uwongo.

Sitaki tudanganyane, ndipo haya nakwambia,
La kweli tuambiane, zitakate zetu nia,
Na ufanyapo mangine, mwisho wakwe ni vibaya,
Vitakuja kutukia, tus'ambiane uwongo.

Vitakuja kutukia, uyaone matukuzi,
Na kwamba wapendelea, n'elezato waziwazi,
Nipate yafahamia, nifunue masikizi,
Tuuwate upuuzi, tus'ambiane uwongo.

8.6. Confronting the Formal Models with Empirical Reality

To ensure the epistemological robustness of the formal models of the three meters (*utenzi*, *wimbo*, and *shairi*), we must remember their dual nature: they are formal frameworks anchored both in the rules stated by the master poets of the Swahili language and in the materiality of the texts. Their rigor is not an a priori construction, but a formalization of what the tradition prescribes and the corpus attests. Nevertheless, any formal matrix—and its corresponding algebraic equation—must be subjected to severe tests against the corpus. A robust theoretical framework is iteratively refined through this continuous confrontation with empirical reality, relying on constant cycles of data modeling and verification.

It is precisely this imperative that closes the formal work of this chapter and opens the empirical investigation of the next. The matrices formalized above—for the *utenzi*, the *wimbo*, and the *shairi*—are the product of repeated confrontation with the manuscripts and with the rules stated by the tradition, but they have not yet been submitted, in this book, to systematic tests stated in advance: at this stage, they are internally consistent, algebraically tractable, and designed to generate the permutations the tradition describes. But internal consistency is not, by itself, proof of anything beyond the model's own coherence. A formal model can be perfectly self-consistent and still fail to describe the object it claims to describe; the history of linguistics offers no shortage of such cautionary examples. The only way to know whether these matrices capture something real about the Swahili line as it is composed and read—rather than merely something convenient about how it can be described—is to confront them with data the models did not generate: the corpus itself, as printed and as read aloud.

Chapter 9 submits the eight conditions of falsifiability set out in Chapter 7—four bearing on the syllabic and rhyme architecture, four bearing on the competing accentual hypotheses—to a symmetrical protocol of verification across a corpus spanning nearly two centuries and all three classical meters: an exhaustive derivation of the accentual structure of every line from its lexicon, an exhaustive auditory screening of formal readings by L1 Swahili speakers, and Praat spectrographic measurement of a sample of lines, reported for information. Two control corpora of ordinary prose provide the baseline. Whatever the outcome, it is this confrontation, and not the formal elegance of either model, that will determine which description, if any, corresponds to the object under study.

Chapter 9. Empirical Proofs: Experimental Verification of the Formal Syllabic and Accentual Models

9.1. Epistemological Prelude: Falsifiability vs. Ptolemaic Epicycles

Before descending into the measurements and empirical datasets of this chapter, we must establish a rigorous epistemological firewall. In the history of science, when a dominant paradigm begins to fail under the accumulating weight of contradictory evidence, its proponents rarely abandon it immediately. Instead, they resort to the construction of "Ptolemaic epicycles"—arbitrary conceptual patches and theoretical contortions designed exclusively to explain away statistical anomalies while keeping a flawed core intact (a conservative practice famously identified as "saving the phenomena," cf. Duhem, 1908/1969).

Just as medieval astronomers added increasingly complex and speculative sub-orbits to save the geocentric model of the universe from the reality of planetary motion, the colonial and contemporary advocates of the accentual school have introduced a myriad of conceptual epicycles to save the exogenous stress-based model of Swahili poetry. When the empirical reality of the Swahili line refuses to yield the symmetrical two-beat rhythm they expect, they do not question the axiom; they invent the "*Grenzschema*" (Vierke, 2011; see Section 2.3.4), the "rhythmic tendency," the caesura as the trace of a foot (Minerba, 2025, p. 298; see Section 5.3.1), or the "irrelevant syllables" (Vierke, 2011; see Section 2.3.3). These are not scientific observations; they are theoretical life-rafts. They conveniently appear only where the exogenous theory needs them to maintain models built on the trochee, taken in its accentual sense, and they disappear the moment the data are examined line by line.

In this chapter, we eradicate these epicycles by applying the strict Popperian criterion of falsifiability (Popper, 1959). A metrical theory that can stretch to accommodate any contradiction is not a scientific theory; it is a dogma. We are not interested in phantom "tendencies"

that hold true only when a specific performer feels musically inspired to stretch a vowel—the ultimate manifestation of the performative fallacy, which fatally confuses the deterministic written grid of Layer 1 with the variable melodic execution of Layer 3 (see Section 3.3)—nor are we interested in a "trochee" foot that exists nowhere in the metrical rules of Swahili poetry as described by the poets themselves (see Chapter 4).

A scientific law in poetics, if it is to be called a law, must be deterministic and measurable. If the generative metrical grid of the *mtunzi* (the composer) is syllabic, the arithmetical count of the *mizani* (syllable count) must hold across the corpus, its only exceptions being the lines that the tradition itself names as defective (*guni*), and these must remain exceptions. If there is a generative metrical accent driving the line, it must manifest not in the subjective imagination of some Western analysts, but in what can be checked independently of them: in the lexical structure of the lines and in what formal readings realize of it, both compared with ordinary prose. The following sections provide the exact experimental protocol to separate the structural law of the poem from the acoustic illusions of its interpreters.

9.2. Protocol of Experimental Reproduction: The Three-Step Verification

In the pursuit of Open Science and total methodological transparency, it is imperative to detail exactly how our empirical experiments can be reproduced, verified, or challenged by any independent researcher. The falsification of the accentual model and the verification of the endogenous syllabic matrix do not rely on subjective literary interpretation; they rely on a deterministic reading of the printed text, a systematic auditory reading of the recordings, and statistical comparison with a prose baseline, with instrumental measurement added on a sample for information.

To reproduce our results and test the formal parameters of the Swahili metrical system, researchers must adhere strictly to the protocol set out in sections 9.2.1 to 9.2.4, whose core is the three-step verification of section 9.2.3.

9.2.1. The Recorded Corpus: Reading vs. Singing (Excluding Layer 3)

The most critical vector of error in the historiography of Swahili metrical studies is the severe methodological confusion between the text's formal foundation—where Layer 1 (the invariant syllabic principles of *vina na mizani*) actively generates Layer 2 (the cognitive and architectural structures ultimately spatialized on the page)—and Layer 3 (the variable, musical realization of the sung performance).

To reproduce our data, the researcher must explicitly record spoken texts, not sung performances. The corpus must consist of formal recitations (reading aloud) of the poetry. The moment the researcher records a *mghani* (singer) applying a melody (*mahadhi*), the data is instantly contaminated by musical isochrony, melismas, and performative rests, shifting the analysis from structural poetics to ethnomusicology. Our experiments exclude Layer 3 by recording a formal reading rather than a sung performance.

9.2.2. Speaker Selection and Acoustic Control

The selection of the informant is paramount to establishing a controlled experimental baseline. The speaker must be an L1 speaker of Swahili, deeply embedded in the universally shared phonological realities of the language. The speaker must possess clear articulatory precision, avoiding excessive colloquial elisions that belong exclusively to rapid spontaneous speech. The instruction given to the speaker is absolute: read the text formally, without attempting to sing or artificially force a poetic rhythm.

Ideally, to achieve absolute methodological perfection, the acoustic dataset would feature native speakers of the specific regional dialects (such as *Kiamu* or *Kimvita*) corresponding to the origin of each classical manuscript. In practice, however, coordinating such a diverse and specialized linguistic panel presents significant logistical challenges. Consequently, with the notable exception of the contemporary composition *Kiswahili!!* by Sheikh Mahmud Ahmed Abdull Kadir (*Ustadh* Mau), which is read aloud by the author himself, the classical corpus is voiced by a single reader: Fatuma Shaaban, a twenty-five-year-old speaker of 21st-century Standard Swahili based in Tabora, Tanzania.

She is a literate reader with a practical understanding of Swahili poetics. This property of the reader must be declared as a limitation rather than an advantage. Her familiarity with Swahili poetics guarantees that she does not stumble over the specialized register, but it also means that she approaches these texts as poetry. A reading by a competent reader of poetry is not a neutral utterance, and the possibility that a recitation habit is carried into what was intended as a plain reading cannot be excluded. This is one of the reasons the accentual argument of this chapter is not made to rest on the reading alone.

Crucially for the acoustic integrity of the analysis, both our L1 Swahili speakers possess full L2 proficiency in Arabic, ensuring the precise phonological realization of the Arabic phonemes frequently embedded within the classical vocabulary.

This same competence has a consequence that must be recorded, because it accounts for a category of exclusion applied in the case studies. A reader fluent in Arabic tends to restore Arabic phonology in the loanwords of the classical lexicon, preserving closed syllables and geminates that the assimilated Swahili form does not have. The result is a line executed with fewer *mizani* than the text contains. These executions are identified by direct comparison with the source text and excluded from the statistical pool, and every such line is listed with both its source form and its executed form.

The combined literary and linguistic competence provides a critical advantage: it guarantees a highly fluid, uninterrupted recitation. Because the readers do not hesitate, stumble, or insert unnatural pauses when encountering specialized poetic vocabulary or the Arabic loanwords frequently embedded in the classical corpus, the continuous flow required for a reliable auditory annotation is preserved.

However, the overarching constraint—namely, relying on an acoustic dataset of only two speakers, with a single modern reader voicing the entirety of the classical corpus across non-native dialectal zones—weighs less than it seems, because the experiment rests on macro-Swahili properties at the synchronic level: the syllable count of a written line, and the lexical stress that the language assigns to its words. The constraint bears only on what a reading realizes of that structure, which is recorded as one reading among others. Across the entire Swahili linguistic continuum, the cognitive and articulatory process of syllabic segmentation operates on a uniform set of morpho-phonological principles. A native speaker of *Kiamu*, *Kimvita*, or

Standard Swahili is perfectly capable of navigating and vocalizing a manuscript originating from a different dialectal zone. Because the recordings are strictly straightforward formal readings rather than spontaneous speech, the lack of native regional dialectal fluency becomes structurally irrelevant. While a reader relying on Standard Swahili competence will naturally neutralize or adapt specific dialectal phonetic and phonological features—such as the precise articulatory strategy used for a syllabic nasal consonant, the phonetic realization of retroflex compounds (e.g., the dialectal /ndr/, which Fatuma Shaaban fluidly translates and voices as the standard /nd/ without any reading latency), or the historical aspiration of certain plosives—this phonetic approximation does not alter the underlying morphological syllabification. Whether a dialectal feature is fully neutralized or merely adapted to standard articulatory habits, the absolute syllable count remains arithmetically intact and identical.

When encountering a specific *Kiamu* form (for instance, reading *zina* "rhymes" instead of the standard *vina*), the modern reader intuitively applies these universal Swahili syllabification rules, executing the requisite metrical units without requiring the intrinsic phonetic habits of a 19th-century coastal speaker. In any case, a literate reader like Fatuma Shaaban is already familiar with these archaic and regional features through her knowledge of Swahili poetics. Many of these Northern forms are conventionally preserved within *Kingozi*[83], the traditional literary and poetic dialect of Swahili.

---

[83] The exact linguistic nature of *Kingozi* has historically been the subject of debate. Early scholarship often viewed it as a purely artificial, archaizing literary jargon rather than a genuine spoken language. Notably, Sacleux (1909) defined *Kingozi* as an exclusive, esoteric poetic register—a "free form" (*forme libre*)—that deliberately borrowed from various Northern dialects, primarily *Kigunya*, *Kiamu*, and *Kimvita*, to achieve a stylized, archaic purism. However, computational linguistic analysis by Klein-Arendt (1987) resolved this dichotomy by demonstrating that *Kingozi* is in fact twofold: it preserves the structural traces of a genuine, now-extinct ancient dialect (genealogically aligned with Northern varieties like *Kiamu*, *Kipate*, and *Kisiu*), while simultaneously functioning as a literary medium overwhelmingly characterized by *Kiamu* features.

Ultimately, both observations confirm that *Kingozi* practically functioned as an interdialectal, supralocal poetic *koiné*. Classical Swahili poets systematically employed this elevated literary language regardless of their geographical origins, freely mixing regionalisms. Consequently, a modern reader literate in the Swahili poetic tradition—such as our informant—is already intimately familiar with this morphological heterodoxy. When encountering these non-standard archaic forms, she intuitively processes them as the conventional morphology of the poetic register itself, entirely bypassing the need for native fluency in any specific historical Northern dialect.

Furthermore, as Wamitila (2006) outlines in the introduction to his *Kamusi ya Ushairi* [Dictionary of Poetry] (see *Utangulizi*, unpaginated), Swahili poetics operates on a fundamental distinction between specialized poetic language (*lugha ya kishairi*) and ordinary everyday speech (*lugha ya kawaida*). Consequently, when a literate reader encounters a specialized form—such as *moo* ("legs") instead of the Standard Swahili *miguu*, or *nyuni* ("bird") in place of *ndege*—they do not experience it as an exposure to an unfamiliar or incomprehensible regional dialect. Rather, they intuitively process it as a recognized morphological convention of the poetic register. This literary competence effectively neutralizes the potential structural obstacle of dialectal divergence during the formal reading process.

Yet, this practical setup establishes a rigorous condition of falsifiability, which is essential to solidify the core scientific demonstration. Because the primary objective of this mapping is to systematically evaluate the exogenous accentual models against the endogenous syllabic matrix, the lexical stress structure of each line must be derivable with confidence. This methodology relies on a fundamental, testable hypothesis: lexical stress, falling on the penultimate syllable of the phonological word, is a macro-Swahili phonological property. By employing the term 'macro-Swahili', this study operates at a level of structural abstraction, identifying the overarching phonological principles that remain stable across the Swahili continuum in both time and space. At this level of abstraction, isolated variations in peripheral dialects are structurally negligible. The historical literary coastal and insular dialects (*Kiamu*, *Kimvita*, *Kiunguja*) and modern Standard Swahili do not merely belong to this overarching system; they provide its absolute empirical validation. Within this shared linguistic core, the penultimate stress rule is stable at the synchronic level across the geographical dialectal spectrum.

Building upon this synchronic uniformity, we project this stability diachronically. The working hypothesis is that the two centuries elapsed since the composition of these older texts do not constitute a sufficient linguistic timeframe for such a deeply anchored phonological property to have mutated. While we can never claim absolute certainty—as no audio archives from the 18th or 19th centuries exist to verify historical pronunciation—the extreme synchronic stability of this rule strongly supports the diachronic assumption. If this foundational premise of diachronic and synchronic stability were proven false, the entire experimental mapping and its subsequent conclusions would be instantly

invalidated. However, operating under this scientifically reasonable postulate, the lexical stress derived by the modern penultimate rule is the closest, most rigorously justifiable approximation of the historical stress patterns available to contemporary research.

This diachronic projection is further supported by comparative Bantu phonology. Penultimate stress and lengthening is not a Swahili-specific idiosyncrasy but a well-documented areal feature of the Eastern and Southern Bantu zones (Hyman 2013). Furthermore, Swahili's development of a predictable stress-accent system, fixed by rule on the penultimate syllable — replacing the inherited Proto-Bantu tonal system retained by most of its sister languages — is understood in the comparative-historical literature as an early, structural consequence of language contact (Nurse & Hinnebusch 1993). A categorical typological shift of this kind, once established, does not fluctuate on a century-to-century basis; it belongs to the same class of stability as other core morphophonological parameters of the language, not to the class of labile surface habits.

By utilizing a single contemporary reader for the classical corpus, the experiment intentionally and systematically homogenizes the oralization process. This aligns with the absolute instruction given to the informant not to sing or chant the text. This methodological decision deliberately sets aside Layer 3 — the indeterminism, the performative variations, and the subjective melodic interpretations inherent to the chanted or sung execution of the poem. The explicit objective is to record a precise, unadorned reading of the text as it stands on the page, which carries the syllabic count and the rhyme boundaries of Layers 1 and 2 and no accentual notation whatever. She executes the macro-Swahili accentual rules of the language while submitting entirely to the formal metrics of the *vina na mizani* system. This standardized output provides an unembellished record of a formal reading. It is precisely this record, set beside the accentual structure derived from the printed lines, that allows us to test — and ultimately to falsify — the hypothesis that classical Swahili verse is governed by a recurring metrical beat rather than by syllabic arithmetic.

To ensure absolute methodological transparency, the raw acoustic data has been deposited in online open-access archives, specifically via Zenodo[84]. These repositories allow readers and researchers to

[84] Zenodo is a multi-disciplinary open research repository developed under the European OpenAIRE program and operated by CERN (the European Organization for

independently verify the accuracy of the syllabic computation and accentual mapping presented in this volume. Furthermore, providing unrestricted access to the raw audio files encourages the scientific community to conduct independent verifications, whether by auditing this primary dataset against the book's analyses or by recruiting their own speakers for comparative testing. The exact DOIs and web locations to access these audio archives are fully detailed in Section 9.2.7 (Open Science Repositories and Methodological Freedom).

9.2.3. Syllabic Segmentation and Accentual Mapping: The Tripartite Funnel

The debate between the endogenous syllabic matrix (*vina na mizani*) and the exogenous accentual paradigm cannot be settled by philological argument alone, nor by acoustic measurement alone. The recurring methodological trap of the accentual school has been to conflate the deterministic lexical accentual structure of the lines of the poem with the variable prosodic surface of their oral realization. To hold these two orders of fact apart, the verification proceeds through a three-step funnel. Each step yields its own dataset, each dataset is published separately, and no step may revise the one before it.

A point of vocabulary must be fixed at the outset, since the whole of this chapter depends on it. Layer 1, as defined in section 3.3.1, is the syllabic law: *vina* and *mizani*, the positional boundaries and the arithmetic enumeration of syllables. It is formally flat and contains no accentual specification whatever. The accentual structure of a line — which syllables carry lexical stress, and how many groups they form — is not a property of Layer 1 but a property of the Swahili language, produced by the fixed penultimate stress of the phonological word. It is precisely the accentual school that mistakes this phonological reflex for a metrical parameter. The funnel below therefore measures the accentual structure not because it belongs to the meter, but because the accentual school claims that it does, and because a claim can only be falsified on its own ground.

The notation used throughout is that established in sections QRG.1 and 2.5: G for the number of accentual groups in a line, A*i* for the *i*-th accentual nucleus, |A*i*| for its syllabic volume, P*i* for the positional index of the stressed syllable anchoring A*i*, the hyphen for a stressed

Nuclear Research), which provides the robust infrastructure necessary to guarantee the long-term preservation of open-science artifacts.

syllable, u for an unstressed one, and the central dot for concatenation. Two instantiations of G are distinguished and carried through the whole chapter: **G-lex**, the number of accentual groups the lexicon of the line contains, derived by rule; and **G-aud**, the number realized in the formal reading, established by ear.

1. Syllabic segmentation and the lexical accentual structure (G-lex)

The first step operates on the lines of the poem, independently of any measurement. Each line is segmented into syllables, *mizani*, and the fixed penultimate stress of the Swahili phonological word is applied to each word of the line, yielding the values of A*i*, |A*i*|, P*i* and G-lex.

Clisis. A monosyllabic token carries no stress of its own and attaches to the word that follows it; a line-final monosyllable attaches to the word that precedes it. Adding a proclitic at the front of a nucleus does not move that nucleus's penultimate, so the stressed syllable of a nucleus A*i* of |A*i*| syllables is always its syllable |A*i*| − 1.

The sequence of nucleus *volumes.* Because the stress is invariably the penultimate of its nucleus, the sequence $|A_1|\cdot|A_2|\cdot\ldots\cdot|AG|$ determines the entire matrix. The sequence 2·2·4 yields $P_1=1$, $P_2=3$, $P_3=7$; the sequence 4·4 yields $P_1=3$, $P_2=7$. Within each nucleus the final two syllables constitute the trochaic core in the sense of section 2.5.1, and the syllables preceding them are pre-nucleus. The notation is therefore not a rival to Vierke's: it is her own bipartite figure, written as a sequence of volumes. It follows that the ideal matrix she calls the *Grenzschema* is the single sequence 4·4, and that her permitted asymmetrical fractionations 2-6, 3-5, 5-3 and 6-2 are the sequences 2·6, 3·5, 5·3 and 6·2.

This step carries the principal argument of the chapter, and it is the only result in this study that depends on no recording, no listener and no software. Any reader who knows the penultimate rule recomputes it from the lines as printed. Its evidential value follows from the accentual school's own claim: if the composer built the eight-syllable measure around a bipartite template, that design must be visible in the morphological material selected to fill the measure. It should be said plainly that a distribution establishes what the lexicon produces and not what the poet privately intended; the falsifiable claim defended here is the former.

The executed lexical form. A critical distinction regarding the dialectal nature of the corpus must be entered, and it is documented line by line in each case study. The mapping of the lexical forms is strictly based on how the reader actually executes them. For the historical classical corpus, this means the mapping relies on the Standard Swahili adaptation rather than the original Kiamu orthography of the source text: where the Tanzanian reader reads *mtu* for *mthu*, *ndiyo* for *ndriyo*, or *yake* for *yakwe*, the executed form is the one mapped. Conversely, for the poem *Kiswahili!!*, read by its own author (Ustadh Mau), no such adaptation occurs; the author naturally executes his own dialectal forms, and the mapping reflects this exact authorial execution. This is required for the two columns to be commensurable. Every divergence between the source text and the execution is listed in a dedicated table, together with its effect, if any, on the syllable count.

2. The auditory realization (G-aud)

A formal reading aloud is not Layer 1 itself. It actualizes the lexical accentuation of the words, and it superimposes a phrase intonation, which belongs to neither Layer 1 nor Layer 3: the reduction described here occurs in ordinary reading, with no melody at all, and is therefore not the melodic overlay (*mahadhi*) of the *mghani*. As the phrase contour is imposed, certain lexical nuclei are realized while others are de-accentuated, or fused into a neighboring nucleus. This surface realization yields G-aud, established by systematic line-by-line listening at reduced speed over the entire corpus.

The two measures are reported side by side for every line, and they address the accentual hypothesis from two different directions. G-lex states what the lexicon of the line provides. G-aud states what a reader actually produces. Because phrase intonation selects among the nuclei available and does not create new ones, G-aud cannot in principle exceed G-lex except by an added prominence, and such additions are counted separately for that reason.

Transitions are counted in both directions. The symbol −P*n* marks a lexical nucleus that is not realized, and +P*n* a realized prominence that the lexicon does not predict. The two counts bear on the accentual thesis in opposite ways. A deletion shows the voice pruning a structure it did not create. An addition, by contrast, would constitute a beat not derivable from the lexicon, which is what a generative accentual meter

requires in order to exist. Every addition found in a corpus is therefore published individually, with its line and its position.

Confidence. Auditory decisions are not uniformly secure, and the insecure ones are marked rather than silently resolved. A question mark attached to a positional index records that the decision at that position is uncertain: P2=5? marks a nucleus that is heard but judged weak and is nonetheless counted, and −P5? marks a lexical nucleus judged unrealized without full confidence. The position is counted exactly as written; the question mark registers the uncertainty and does not resolve it. No such mark appears in the lexical column, which is derived by rule and admits none.

Two consequences follow, and both are reported with each corpus. The number of uncertain positions is itself a measure of how clean the annotation is, and it is published. And for the lines that carry one, G-aud is reported as an interval rather than as a point: the lower bound obtained by discarding every uncertain nucleus and confirming every uncertain deletion, the upper bound obtained by doing the reverse. The corpus statistics are then given under both bounds, so that a reader may see immediately whether the result depends on the doubtful cases. The marks are assigned before any instrumental measurement is made, and the lines carrying them are natural candidates for the sample described in step 3 below, where the instrument can report on a decision the ear did not settle without revising it.

Exclusions. Two categories of line are removed from the statistical pool, on criteria fixed before counting. Arabized executions preserve closed syllables and geminates and thereby displace the stress positions mechanically. *Guni* of reading are lines in which the reader, working on a nineteenth-century Kiamu text she does not practice, adds or omits a syllable that the source text does not contain: the deviation originates in the execution and not in the meter. The distinction between the two categories and the genuine *guni* of the source text is settled by direct comparison with the text supplied to the reader, not by impression. Genuine *guni* of the text are retained, including where they contradict the syllabic constant that this study defends.

Intra-annotator control. The annotation is the work of a single listener, and the corpus cannot be double-annotated within the constraints of independent research. A blind re-annotation of a subset was therefore carried out one week after the first pass, without reference to the original table. The agreement rate and, more

instructive still, the direction of the disagreements are reported with each corpus. This procedure measures the reproducibility of a single judgment and not its validity: a systematic bias would reproduce itself across both passes and would go undetected by it.

3. Instrumental verification on a hostile sample (G-inst)

The exhaustive verification of the corpus is auditory. Instrumental analysis, conducted in Praat through the Parselmouth library, is applied to a small sample of lines and to no more than that. This is a deliberate division of labor and not a shortcut: the demonstration of steps 1 and 2 rests on an exhaustive count, and the function of the acoustic work is to establish that the categories being counted have measurable physical correlates. No statistic in this chapter rests on the instrumental step, and none is claimed to.

Because a sample chosen by the analyst invites the objection of selection, its composition is fixed before any result is known, and it deliberately includes lines that favor the adverse model:

- lines conforming exactly to the *Grenzschema*, so that the accentual model is tested on its most favorable ground;
- asymmetrical lines yielding G-aud = 2, where the count agrees with the accentual model but the positions do not;
- lines yielding G-aud = 3 or 4;
- a line carrying an addition, that is, a prominence the lexicon does not predict;
- a line carrying a nucleus flagged as weak;
- two lines drawn at random.

The random draw. A published seed is not by itself a guarantee, since the analyst publishes the seed he chose to keep. The seed is therefore not chosen at all: it is derived from a public string fixed long before the analysis was undertaken, namely the ISBN of this book.

```
seed = SHA-256("978-1-291-70670-3") mod 2³² = 3 946 069 218
population: the retained lines of the corpus, ordered by stanza and
then by line, indexed from 1
draw: random.Random(seed).sample(range(size), size), Python 3.12
```

Only one seed is possible, and any reader recomputes it in a single line of code. Lines already belonging to the reasoned sample are skipped and the next line in the generated sequence is taken; every such substitution is recorded. The draw is performed once, and the lines obtained are analyzed and published whatever their profile.

For each line in the sample, three parameters are measured syllable by syllable — fundamental frequency in semitones relative to the speaker's median and corrected for the declination internal to the line, duration, and intensity — and all three are reported in full, whether or not they agree with the annotation.

Rule of non-revision. Step 3 never revises step 2, and step 2 never revises step 1. The three datasets belong to distinct orders of fact: G-lex is the accentual structure the language assigns to the words of the line, G-aud its realization by a reader, and the acoustic measurements the physical signal itself. Where they diverge, the divergence is published: a nucleus annotated by ear with no measurable correlate is reported as such, exactly as is a measurable prominence the ear did not annotate. The outcome of this step is a concordance rate reported together with its failures, and not a claim that the annotation has been validated. Were the annotation revised to match the measurements, the acoustic step could no longer disconfirm anything, and the protocol would become circular.

The audio recordings of the formal readings are archived on Zenodo, and the scripts that generate every acoustic figure reproduced in this book are available in the computational repository, together with the Docker environment that preserves them against future software drift; all three are described in section 9.2.7. One script corresponds to each line, or each hemistich, illustrated in the book, and is named for the poem, the stanza and the line or hemistich it analyses. The reader may therefore reproduce any figure, alter its parameters, or apply the same procedure to any other line of the corpus. It bears repeating that the derivation of G-lex requires none of this: it is recomputed from the printed lines by the rule stated above, without recording, listener or software.

9.2.4. Epistemological Parameters: Defining the Conditions of Falsifiability for the Competing Metrical Models

The competing metrical models are tested directly against the definitive structures established by the first two steps of our protocol: the syllabification and lexical accentual structure derived from the text, and the syllabification and accentual structure recorded through exhaustive manual auditory screening. Once these syllabic and accentual structures are mapped across the corpus, they are evaluated against the formalized equations of both theoretical frameworks. The

falsification of these models does not rely on the instrumental acoustic extraction. This instrumental reading is explicitly excluded from the falsification conditions; it serves strictly as a transparent physical verification of the acoustic signal for a subset of lines to illustrate how the raw physical parameters behave in reality.

In strict accordance with Karl Popper's epistemology, our methodology establishes neutral, absolute conditions of falsifiability derived directly from the theoretical reduction of both the endogenous *vina na mizani* syllabic system and the exogenous accentual models. Scientific objectivity demands that no theoretical paradigm is immune to empirical refutation. Therefore, if the exogenous accentual models (as formalized in Chapter 2, Sections 2.5.1 and 2.5.2) cannot survive testing against the established lexical and auditory accentual structures of the line, they must be discarded. Equally, if the formalization of the endogenous syllabic principles (as formalized in the Quick Reference Guide and Chapter 8) fails to predict the absolute boundaries of the established syllabification, it too must be rejected.

1. Testing the Endogenous Syllabic Principles (Vina na Mizani): *Formal Matrices and Algebraic Equations*

Condition 1: The Absolute Mathematical Verification of the Syllabic Constants (N, M)

A foundational premise of the *vina na mizani* formalization is that the syllable count is not a musical tendency but a rigid algebraic constant. The condition is tested on the line as printed and on the line as executed, and it asks whether each structural segment contains exactly the required number of syllables. The equation must predict the structural boundaries of the text using strictly natural Swahili phonotactics, without appeal to any performative device.

The treatment of deviations is fixed here, before any figure is given. Deviations are of two kinds and they are counted separately. A deviation produced by the execution — a syllable added or omitted by the reader with respect to the source text supplied to her, or an Arabized rendering that restores closed syllables absent from the Swahili form — is a fact about the reading and is excluded from the pool, with the source line and the executed line both printed so that the decision can be audited. A deviation present in the source text itself is a fact about the poem and is retained in the pool, where it counts against the matrix.

The second kind bears directly on the condition, and the criterion is this. The Swahili tradition does not claim that every attested line satisfies the measure; it claims that a line failing the measure is defective, names it *guni*, and rejects it (Abedi, 1954, p. 16; Robert, 1972). The matrix therefore predicts the norm, and its exceptions are predicted to be exactly those the tradition identifies as faulty. The condition is satisfied if the deviant lines are few and are *guni* in the endogenous sense. It is falsified if deviation is common, or if deviant lines are metrically indistinguishable from regular ones and pass without the tradition marking them — for that would show the syllabic constant to be a tendency of the analyst's making rather than a rule of the system.

Condition 2: The Spatial Locking of the Algebraic Rhyme Variables (a, b vs. X, Y)

The generative equations do not merely count syllables; they mathematically lock specific phonetic sounds (*vina*) to specific spatial coordinates using local variables (e.g., a, b) and global constants (e.g., X, Y). The matrices dictate that local variables must mutate across stanzas, while the global constant (X) must remain completely immutable at the absolute terminal boundary of every stanza across the entire poem. If an empirical analysis of the manuscripts reveals that the terminal rhyme X fluctuates randomly, or that the internal rhymes a and b fail to achieve phonetic homophony at the internal caesuras (+) or line boundaries (//); or if the poet relies on pure semantic assonance rather than strict syllabic homophony, the formal model is violated. A matrix claims deterministic topological engineering; if the rhymes drift outside their assigned coordinates, the equation is empirically falsified.

Condition 3: The Syntactic Impermeability of the Metrical Caesura (+)

In the formalized matrices for the *shairi* and *wimbo* meters, the plus symbol marks the internal caesura, severing the line into hemistiches. The condition asks whether the natural syntax of the text respects that boundary. If the composer frequently forces a single morphological word to bleed across the caesura, placing its prefix in one hemistich and its root in the other, the divided matrix collapses and the caesura is exposed as a typographical convention rather than a compositional constraint.

Condition 4: The Falsification of Topological Permutations and Stanzaic Chaining

For highly complex meters like the 16-syllable *shairi*, the algebraic equation strictly limits the fourth line (L_4) to a finite set of structural permutations defined by the boolean operator | (e.g., $([8,a]+[8,b]) \mid ([8,b]+[8,X]) \mid ([8,b]+[8,a]) \mid ([8,X]+[8,Y])$). If the text generates a final line utilizing a completely random rhyme scheme that falls outside these defined algebraic operators, the overarching generative formula is falsified. Furthermore, for chained matrices like the *Mashairi ya Vidato*, the equation,

$$P = \sum_{i=1}^{k} \left[3\left(8_{r_{i-1}} + 8_{r_i}//\right) \cdot \left(8_{r_i} + 8_X//\right)\right]$$

dictates that the spatial propagation of the rhyme is formally recursive. If empirical observation of the text shows that Stanza k fails to inherit the exact rhyme r_{i-1} from Stanza $k-1$ in the exact hemistich position predicted by the sequence, the topological matrix is falsified. The model must successfully compile the entire poem without a single broken link; otherwise, its formalization is empirically invalidated.

2. Testing the Exogenous Accentual Theory: Rhythmic Axioms and Cognitive Prototypes

The following four empirical conditions (Conditions 5 through 8) directly synthesize the epistemological parameters previously established during the formalization of Clarissa Vierke's baseline model (Section 2.5.1) and Emiliano Minerba's extended framework (Section 2.5.2). While the theoretical justification for these falsification criteria was detailed in Chapter 2—where their respective theories were reduced to testable formal models—they are recalled and centralized here under a unified numbering system for practical application. By aligning them directly alongside the four endogenous syllabic principles (Conditions 1 through 4), we establish a symmetrical, eight-point testing grid that will be systematically applied to the lexical and auditory readings throughout this chapter.

A word is required on the standing of the two halves of this grid, since they do not perform the same operation. The accentual models of Conditions 5 to 8 were formulated independently of the corpora tested here and imported into Swahili from other traditions; they can fail outright, and if they do they must be discarded. The syllabic

matrices of Conditions 1 to 4 were formalized from the descriptions of the poets themselves and from the manuscripts, and refined by successive contact with the corpus. They too can fail: a line that contradicts the matrix obliges the matrix to change, and the model advances by this round trip between formalization and verification. Had the *utenzi* matrix been established on a six-syllable definition and the corpus returned eight, the parameter would have been corrected, not defended. The symmetry of the grid lies in the fact that both sets of conditions are stated before the data are examined; it does not lie in the two operations being identical, and this study does not present them as such.

Condition 5: Testing the Bipartite Accentual Axiom (G=2)

The accentual theory posits that the eight-syllable segment is inherently bipartite, containing exactly two accentual groups. The claim is made in two forms in the literature, and the two are tested separately because they are not refuted by the same observation.

As an axiom, G = 2 admits no exception, and a single line whose lexicon yields a different number of accentual groups falsifies it. This is tested on G-lex, derived by rule from the printed line and requiring no recording, and again on G-aud. The proportion of lines with G ≠ 2 is reported on both measures; it does not affect the verdict on the axiom, which turns on the existence of counter-examples, but it measures how far the configuration is from being even a dominant tendency.

As a cognitive prototype — the term Vierke herself uses — G = 2 is a statistical claim, and a statistical claim requires a baseline that is not itself derived from poetry. That baseline is supplied by the prose control corpus described in section 9.2.5: eighty lines of ordinary Swahili prose, each of exactly eight *mizani*, composed with no rhyme, no stanzaic grouping and no thematic continuity, analyzed by the same rule and read by the same speaker. If the distribution of G in classical verse is indistinguishable from the distribution obtained in prose, then the bipartite configuration is a property of Swahili word structure inside an eight-syllable frame and not a metrical design, and the prototype is falsified as a generative principle. If the poetic distribution departs from the prose baseline, the accentual school has a genuine phenomenon and this study will report it as such.

Condition 6: Testing the *Grenzschema* as a Cognitive Prototype

If the accentualist theory (Vierke, Section 2.5.1, and Minerba, Section 2.5.2) posits a cognitive prototype locked at positions 3 and 7 ($P_1 = 3, P_2 = 7$), what is the statistical frequency of this exact alignment within our corpus?

The purported rhythm must exert a statistically significant gravitational pull on the text to be considered an active generative intent.

We evaluate this claim on both measures and at two degrees of strictness, since the accentual school states it at two. In its strict form the *Grenzschema* is the single sequence $|A_1| = 4$, $|A_2| = 4$, with $P_1 = 3$ and $P_2 = 7$. Its frequency is established first on the lexicon of the printed lines, where it depends on no recording and no listener and may be recomputed by any reader, and second on the realization.

In its extended form the criterion is the parity of the first stressed position. Minerba admits as trochaic any disposition placing the stresses on odd-numbered syllables, which is to say $P_1 \in \{1, 3, 5\}$, and classes $P_1 \in \{2, 4\}$ as permitted variations that do not produce the trochaic impression (Minerba, 2025, p. 220). The proportion of lines falling on either side of that division is therefore reported for every corpus.

Both proportions are compared with the corresponding proportions in the prose control corpus. The prototype must exert a measurable pull on the text to count as an active generative intent: if it occurs no more often in classical verse than in ordinary prose confined to the same syllabic frame, it is not a metrical prototype but the arithmetic of Swahili word length, and it is falsified as a foundational generative rule. The extended criterion is falsified by a different observation: if the lines that satisfy it are not the majority of those that satisfy the bipartite axiom, the rhythm is not what the axiom describes.

Condition 7: Testing the Periodicity of Asymmetrical Distributions ($|A_i|$)

A metrical foot is defined by recurrence. It is therefore not sufficient for a line to carry a given number of accentual groups; the groups must recur at a stable interval, within the line and across lines. The condition is tested at both scales, and each has an explicit criterion.

Within the line, the test requires no statistics and no recording, and it bears on the nucleus rather than on the accent. A foot is a unit that repeats; a line therefore exhibits an internal periodicity only if its accentual nuclei are of equal volume, that is, if $|A_1| = |A_2| = \ldots = |AG|$. The sequence 4·4 satisfies this, and so does 2·2·2·2; the sequences 3·5 and 2·2·4 do not, since they juxtapose units of different sizes, which no metrical system would call feet. The proportion of lines whose nuclei are all of equal volume is therefore reported for each corpus. This test is decisive on its own, because it holds even if the bipartite axiom of Condition 5 were granted in full: a line may carry exactly two accents and still contain no repeating unit whatever.

Across lines, the test asks whether the poet sequences these configurations deliberately, and the criterion is independence: the observed frequency of identical sequences of nucleus volumes in consecutive lines is compared with the frequency expected if each line's configuration were drawn independently, given the distribution of configurations in the corpus. If the observed frequency does not exceed the expected one, the poet is not sequencing anything, and the succession of accentual profiles is governed solely by the morphological size of the words selected to fill the syllabic count.

Condition 8: Testing the Tetrasyllabic Foot against Extreme Metrical Boundaries

Minerba derives the classical Swahili line from an imported tetrasyllabic template and states the derivation for the three meters examined in this study:

> The major innovation in the history of the *utenzi*, *shairi* and *wimbo* metres is their adoption of trochaic rhythms as constitutive elements of their abstract patterns. This is evident, historically, in their progressive adoption of syllabic lengths of multiples of four. (Minerba, 2025, p. 217)

The claim is therefore general, and the condition tests it against the boundaries the corpora impose. Two failures are distinguished, and they are not of the same kind.

The first is arithmetical and requires no data beyond the matrix. A segment of two or three syllables cannot host a four-syllable foot, and a hyper-fractionated configuration such as 6 + 3 + 4 + 2 cannot be decomposed into tetrasyllabic units by any partition. Where the corpus presents such segments, the derivation fails on formal

grounds, before any question of prominence arises. This is the ground on which the *wimbo* is decisive and the *utenzi*, whose line is a multiple of four, is not.

The second is empirical and applies where the arithmetic permits the template. It asks whether the composer takes it up: the proportion of lines containing a tetrasyllabic nucleus, and the proportion partitioned entirely into such nuclei, are reported for every corpus.

One further claim is tested separately, because Minerba states it without qualification:

> [...] the four syllable hemistich of [sic] can contain either a phonetic word of 4 syllables (such as, for example *amekuja*, *na wazazi*) or two words of two syllables each (*watu wote*): the stresses are therefore always on odd syllables and the trochaic rhythm is necessarily reproduced. (Minerba, 2025, p. 221)

The enumeration is exhaustive and the conclusion follows. A segment of four syllables admits only these two partitions once monosyllabic formatives are treated as proclitic, and both place the stress on an odd position. What the claim establishes, however, is narrower than what it is invoked to support. The stress falls where it does because a four-syllable segment leaves the penultimate rule no other position to occupy, not because a trochaic template has been imported and applied. The configuration is a consequence of the segment length and of Swahili word structure, and it would obtain identically in prose confined to four syllables.

9.2.5. Corpus and Selection Criteria

To resolve the metrical debate, the experimental verification must be entirely immunized against any suspicion of selection bias. The corpus subjected to formal and empirical verification in this chapter was constructed according to strict, objective epistemological criteria to ensure the most faithful and representative results possible. By establishing this uncompromising methodological foundation, any stark contrasts resulting from the falsification protocol—such as a dominant, robust verification rate for one model paired with near-total disqualification for another—are established as objective structural realities rather than the byproduct of a tailored dataset.

1. List of Selected Works for the Corpus

The literary works selected for analysis, and their recordings, are drawn from two distinct open-access Zenodo repositories. The first five classical texts correspond exactly to the numerical sequence of the audio files archived in the Fatuma Shaaban corpus. The final contemporary composition is archived independently in its own dedicated repository. The corpus comprises the following works:

The Fatuma Shaaban Classical Corpus (Zenodo)

1. ***Utenzi wa Rasi 'lGhuli*** (*utenzi* meter). Composed circa 1850 in Bagamoyo by Mwalimu Mgeni bin Faqihi. This epic, containing 4,584 stanzas, is the second-longest known Swahili epic and was historically performed at communal events. The mapping covers the first 40 stanzas (160 octosyllabic lines) of the edition. (Mgeni bin Faqihi, 1979).
2. ***Utendi wa Haudaji*** (*utenzi* meter). Composed in Lamu during the 19th century by an unknown author. This text serves as the foundational corpus utilized by the exogenous school to establish the *"Grenzschema"* accentual theory. The mapping covers the first 40 stanzas (160 octosyllabic lines), on which this accentual model is tested statistically. (Vierke, 2011).
3. ***Nyimbo*** (*wimbo* meter). This track aggregates three distinct works, demonstrating the structural versatility of the *wimbo* meter. The mapping covers each poem or fragment in its entirety:
    - 3a. ***Wimbo wa Miti*** (*wimbo* meter, 2+8 rupture matrix). Composed circa 1800–1850 in Lamu by an anonymous female poet. This allegorical riddle (*zifungo*) was collected in 1989 from the scholar Zaharia Maimun. (Mbele, 1996).
    - 3b. ***Sina Hali*** (*wimbo* meter, 4+8 asymmetrical matrix). Anthologized as example "b" to explicitly illustrate bicolumnar asymmetry on a strict tetrasyllabic baseline. (Shariff, 1988).
    - 3c. ***Kukwepuka nana*** (*wimbo* meter, 6+3+4+2 hyper-fractionated matrix). Composed by the master poet Sheikh Ahmed Sheikh Nabhany (1927–2017) and anthologized as example "k". (Shariff, 1988).

4. ***Dua ya Kuombea Mvua (Waji** Waji*) (*shairi* meter, 6+9 *takhmisa* matrix). Composed by Sheikh Muhyddin bin Sheikh bin Jahathwan Al-Waily (1798–1870). The mapping covers the complete 14-stanza structure (140 hemistiches) of the prayer. (Chiraghdin & Nabhany, 1987).
5. ***Uwongo*** (*shairi* meter, 8+8 *vidato* matrix). Composed by the master poet Muyaka bin Haji al-Ghassaniy (1776–1840). The mapping covers the entire 4-stanza structure (32 hemistiches), including its stanzaic chaining. (Hichens, 1940).
6. **The Prose Control Corpus.** To the six literary works is added a seventh set of materials which is not a work and is not part of the poetic corpus: eighty lines of ordinary Swahili prose, each of exactly eight *mizani*, composed for this study with no rhyme, no stanzaic grouping and no continuity of subject between lines, covering all thirteen possible partitions of an eight-syllable line into accentual groups. It was read aloud by the same speaker under the same instructions and is deposited with the rest of the data. Its function is to establish what the Swahili lexicon produces inside an eight-syllable frame when no metrical intention whatever is present, and thereby to supply the baseline that Conditions 5, 6 and 7 require. Without such a baseline, a statement that a configuration occurs in classical verse more often than chance has no measurable content. An eighth set completes it: twenty further sentences of ordinary prose, under the same eight-syllable constraint but with the marked word orders characteristic of the classical corpus, read by the same speaker under the same instructions (section 9.8). They are compared with the prose baseline and never merged into it.

The Ustadh Mau Contemporary Corpus (Zenodo)

- ***Kiswahili**!!* (*shairi* meter, contemporary composition). Composed in Kiamu by Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau) in 2003, and read aloud by the poet in Lamu in September-October 2005, from a printed sheet in Roman script, during the present author's fieldwork (Abdulkadir & Frankl, 2013, p. 4; Roy, 2013a, p. 10). The mapping spans the entire 18-stanza poem (144 hemistiches) and is derived from that recorded reading, archived independently (Roy & ibn Kadir, 2016).

2. Corpus Selection Criteria

Criterion 1: Diachronic Persistence and Canonical Authority

Much like the rules of classical music which, once formalized, continue to strictly dictate composition and performance centuries later, this corpus tests the historical persistence of the Swahili formal matrix. Spanning from classical 19th-century masters (such as Muyaka bin Haji al-Ghassaniy's *Uwongo* and Mgeni bin Faqihi's *Utenzi wa Rasi 'lGhuli*) to contemporary composers (Sheikh Nabhany, Ustadh Mau), the selection proves that this arithmetic and syllabic grid is transmitted through time without structural alteration.

Criterion 2: The Epistemological Necessity of the "Adverse Terrain"

To ensure absolute scientific fairness, the exogenous models must be tested on the exact texts utilized to formulate them. The *Utendi wa Haudaji* was explicitly selected because it is the foundational text upon which Clarissa Vierke built the modern accentual school. Furthermore, the 4+8 asymmetrical *wimbo* fragment was deliberately chosen because its segment lengths are perfect multiples of 4, providing the optimal arithmetic environment for Emiliano Minerba's tetrasyllabic metrical foot hypothesis.

Criterion 3: Typological Exhaustivity

As strictly mandated by the fundamental scope of this research—a data-driven manifesto for classical Swahili meters—the dataset must exhaustively cover the generative pillars of the tradition. The corpus therefore systematically tests the competing formal models across the *utenzi*, *wimbo*, and *shairi* meters.

Criterion 4: Geographic Transversality

Sampling across different nodes of the Swahili sphere (from Lamu to Mombasa and Bagamoyo) guarantees a high-quality, representative dataset. This prevents the methodological risk of testing a localized sub-current. By ensuring this geographical diversity, any conclusions drawn from the data—whether they ultimately validate the accentual model or confirm the endogenous syllabic framework—can be reliably established as pan-Swahili structural realities rather than regional anomalies.

Criterion 5: Falsification by Increasing Complexity (Extreme Boundaries)

A robust generative theory must survive limit cases. The corpus deploys a protocol of falsification by increasing structural complexity. It moves from standard symmetrical lines (8+8) to extreme geometric boundaries: asymmetrical configurations and an increased number of lines (6+8, five lines in *Waji Waji*), hyper-fractionated matrices (6+3+4+2 by Nabhany), and macroscopic stanzaic chaining *(vidato)*. This deliberate structural escalation is designed to observe how both models respond to arithmetic asymmetry and stanzaic expansion. It will rigorously reveal whether the accentual framework and the formal syllabic grid can maintain their operational integrity under extreme geometric constraints, or if either system collapses when departing from standard symmetry.

Criterion 6: Socio-Pragmatic Plurality

To determine whether metrical rules are universal or context-dependent, the dataset must encompass diverse social and performative functions. By including a religious epic alongside the *Wimbo wa Miti*—an allegorical social riddle composed by an anonymous woman—the corpus reflects the broad sociolinguistic reality of Swahili poetic production. This selection is explicitly designed to test whether the competing formal models operate identically across a wide pragmatic spectrum, verifying that the generative rules remain constant regardless of the author's social profile or whether the text functions as a sacred invocation or a secular community challenge.

Criterion 7: Heterogeneity of Sources (The Anti-Editorial Bias Shield)

The dataset is constructed from a diversity of transmission media to neutralize any suspicion of editorial filtering. By utilizing manuscripts reconstructed by Western academia (Vierke), local Swahili anthologies (*Diwani ya Muyaka, Malenga wa Karne Moja*), field collections of oral poetry (Mbele), and a recorded reading by the poem's own author (Ustadh Mau), the data is shielded against both transmission anomalies and exogenous textual bias.

9.2.6. Methodological Limitations and the Horizon of Open Science

While the criteria established above guarantee the epistemological validity of the falsification protocol, intellectual honesty requires acknowledging the current logistical and material boundaries of this dataset. These limitations do not undermine the empirical testing of the competing formal models, but rather define the necessary trajectory for future data-driven research in Swahili poetics.

Limitation 1: Sample Size and the Big Data Imperative

The empirical verification conducted in this study operates on a representative sampling model rather than exhaustive macro-analysis. For instance, in the case of the *Utenzi wa Rasi 'lGhuli*, the mapping is applied to a continuous block of 40 stanzas (over five minutes of recorded reading). However, the complete epic consists of 4,584 stanzas. Logistically, executing a manual auditory mapping and transcribing such a volume of text exceeds the capacity of a single researcher. Furthermore, classical masterpieces such as the widely known *Utendi wa Mwana Kupona*, or the entirety of the *Malenga wa Karne Moja* anthology, remain untested in this specific study. The ultimate horizon for this methodology is a transition to "Big Data" poetics. This requires massive external collaborations to digitize, record, and process the entirety of the Swahili corpus.

Limitation 2: Editorial Dependency and the Ajami Archival Deficit

To test the formal models, this study inevitably relies on Latin-script transcriptions and modern editorial anthologies. For the *wimbo* meter, for example, the field remains heavily dependent on the specific examples provided by Ibrahim Noor Shariff. This highlights a severe data deficit in Swahili studies: we lack digitized, widely available primary sources for certain meters. This study does not test original Swahili-Arabic (*Ajami*) manuscripts directly. This absence is primarily due to the colonial legacy of archival displacement; the vast majority of these original *Ajami* manuscripts were historically extracted and are currently locked in European physical archives. To state that these collections are not yet completely digitized is a profound understatement. Accessing them currently requires prohibitive physical travel and institutional clearance. The reliance on

Latin-script transcriptions highlights the urgent need for a massive, open-access digitization of *Ajami* primary sources, with the hope that these physical archives will one day truly open their doors to the global scientific community through comprehensive digital preservation.

Limitation 3: Geographic Scope and Spatial Transmission

This corpus effectively tests the diachronic transmission of the classical matrix—proving its persistence through time from 19th-century masters to contemporary poets like Ustadh Mau. However, it remains geographically anchored to the coastal and insular Swahili sphere (Lamu, Mombasa, Bagamoyo). Future research must test the *spatial* transmission of this cognitive grid. What happens to the strict formal architecture of *vina na mizani* as it travels deep into the continental interior? Extending the testing to authors from the Lake Victoria region (such as those compiled in the anthology *Malenga wa Ziwa Kuu* by Wallah Bin Wallah, 1988) will be a key next step in measuring the resilience of the macro-Swahili metric matrix across varying geographic distances and diverse Bantu linguistic substrata.

Limitation 4: The Ethnomusicological Horizon and the Exploration of Layer 3

To rigorously isolate the generative metrical rules and avoid the "performative fallacy," this study focused on Layer 1 (the fundamental specifications of *vina na mizani*), on Layer 2 (the resulting graphic and architectural structures on the page), and on the lexical stress that the language assigns to the words, which lies outside all three layers. By requiring the poems to be formally read aloud, we deliberately bypassed the musical distortions inherent to Layer 3 (the oral performance, whether sung, chanted, or cantillated). However, this methodological exclusion does not invalidate the performative dimension of Swahili poetry; rather, it correctly assigns it to its proper discipline. A key future direction for Swahili studies is a dedicated ethnomusicological expansion to fully explore Layer 3. It is necessary to systematically record, catalog, and analyze the vast diversity of traditional melodies (*mahadhi*), cantillation techniques, and the specific endogenous terminologies used to categorize these oral executions. Understanding exactly how the non-deterministic musicality of Layer 3 interacts with the syllabic grid of Layer 1 and with the lexical stress of the words, stretching the first and able to

override the second, will ultimately provide a comprehensive, multi-disciplinary understanding of the complete Swahili poetic experience.

Limitation 5: The Single Annotator

The auditory annotation of G-aud across the whole corpus is the work of one listener, who is also the author of the argument it is used to test. Independent double annotation was not available within the constraints of independent research, and no claim of inter-annotator reliability is made. Four provisions limit the consequences. The principal result of this chapter, the derivation of G-lex, does not depend on the annotation at all and can be recomputed from the printed text by any reader. A blind re-annotation of a subset was carried out at a remove from the first pass, and the agreement rate and the direction of the disagreements are reported. The instrumental sample was composed before any result was known and includes lines favorable to the adverse model together with lines drawn at random on a seed derived from the book's ISBN. And the rule of non-revision prevents the annotation from being adjusted to the measurements. None of this substitutes for a second annotator, and the raw audio is deposited so that any reader may supply one.

9.2.7. Open Science Repositories and Methodological Freedom

In strict adherence to the principles of Open Science, all empirical data and analytical tools utilized in this study have been made publicly available to ensure total reproducibility. The reader is actively encouraged to independently audit and verify the findings presented in this volume. To facilitate this, the data and code are distributed across open-access platforms.

1. The Zenodo Data Repositories (Audio Files)

The complete primary dataset is securely stored across two distinct repositories on Zenodo archiving the audio files. These archives serve as the definitive empirical baseline:

Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau) (Contemporary composition)

Available at: https://zenodo.org/records/19827932 (Accessed: August 29, 2026)

DOI: 10.13140/RG.2.1.3695.1922

IPA transcription and French translation:
https://zenodo.org/records/19827635

Fatuma Shaaban (Classical corpus)

Available at: https://zenodo.org/records/22162387

DOI: 10.5281/zenodo.22162387

2. The GitHub Computational Repository and Reproducibility Protocol

To bridge the gap between the raw audio data and the specific visual evidence presented in the case studies, a supplementary computational repository is hosted on GitHub at:

https://github.com/Mawazoni/Ushairi

It is epistemologically crucial to clarify that the primary methodology of this study relied first on rigorous auditory evaluation (by ear), followed by manual phonetic boundary extraction within the standard Praat software. The computational tools provided in this GitHub repository are not the original mode of discovery; rather, they constitute a dedicated framework engineered specifically for the reader to reproduce the visual evidence (the acoustic curves) utilized to illustrate the stanza examples in this book.

This repository provides:

The complete audio files utilized for the book's case studies. The precise temporal boundaries required to isolate and analyze the specific stanzas are embedded directly within the code of the scripts themselves.

Custom Python Scripts and Parselmouth: To allow readers to recreate the visual data, we provide ready-to-use scripts written in Python (a widely used programming language). For non-computer scientists, these scripts utilize a specialized library called **Parselmouth**. Parselmouth acts as a direct bridge: it allows our Python code to directly interface with and control the core acoustic engine of Praat in the background, entirely bypassing its traditional Graphical User Interface (GUI). By simply running these scripts, the reader can instantly and automatically generate the exact acoustic curves shown in this volume and their detailed descriptions.

3. The Permanent Zenodo Archive and Heavy Assets (GitHub clone and Docker image)

To guarantee absolute long-term accessibility and bypass GitHub's file size limitations, the entire computational repository is fully cloned and permanently archived on Zenodo at:

https://zenodo.org/records/23089282

DOI: 10.5281/zenodo.23089282

Crucially, in addition to the scripts and audio data, this Zenodo record hosts the fully built Docker image, which is too massive to be stored on GitHub.

Docker Containers for Long-Term Preservation: Software environments evolve continuously. In a few years, updates to Python or Parselmouth could render these scripts obsolete, effectively breaking the reproducibility chain. To prevent this technological decay, the critical components of this repository have been frozen inside Docker containers. A Docker container can be understood as a virtualized micro-environment that permanently locks in the exact, specific versions of the software used today. Readers can initially attempt to run the standard Python scripts directly; however, if these scripts fail to execute in the future due to system updates, the reader can simply launch the Docker containers (following the commands provided in the repository) to perfectly recreate the original computational environment of 2026.

4. Absolute Methodological Freedom

While the GitHub repository and its Python/Docker tools serve as a highly efficient, ready-to-use reproducibility toolkit, the reader retains absolute methodological freedom. The provision of this code is not intended to prescribe a single method of verification. Researchers are completely free to ignore the Python scripts entirely, download the raw audio files directly from the Zenodo repositories, and apply their own preferred analytical protocols. The acoustic data can be scrutinized through traditional auditory analysis by ear, processed manually within the standard Praat GUI, or analyzed using entirely different phonetic software environments. Indeed, the author's own methodology for this study triangulated all three approaches to ensure absolute rigor: meticulous auditory listening, manual phonetic boundary extraction within the standard Praat GUI,

and Python-driven computational verification. The ultimate goal is total transparency: the data is entirely open, and the reader is free to interrogate it through the method of their choosing.

9.3. Methodological Scope: Formal Reading vs. Sung Performance

A distinction must be made regarding the acoustic data studied in this chapter. The protocol of section 9.2.3 works on formal recitation, reading aloud, and not on sung performance. Before explaining why, it must be stated what a reading is not. A reading is not one of the three layers, and it gives no access to Layers 1 and 2: the syllable count, the rhyme boundaries and the architecture they build are read on the printed page, where the composer fixed them, and no recording adds anything to them. Nor is stress a property of any layer. It is a property of the Swahili language, which assigns a stress to the penultimate syllable of every phonological word, and the accentual structure that results (G-lex) is derived from the printed line by rule.

What a reading adds is a realization of that structure, and the realization already departs from it. Under the intonation of the phrase, some lexical stresses are realized and others are not, and the pattern of these reductions is not predicted by the syntax (section 3.3.3). The auditory structure (G-aud) is therefore not the lexical structure (G-lex), and the instrumental measurement (G-inst) departs further still, since it records acoustic prominences that the ear may not retain. These are three different objects, and this chapter keeps them apart.

Singing departs from the lexical structure far more than reading does. It is not deterministic: a single syllabic grid supports an unlimited variety of musical settings, from flat liturgical cantillation to highly syncopated performance, and the singer (*mghani*) occupies a different functional role from the composer (*mtunzi*). To satisfy the demands of musical time, singing reshapes the realization of the line through four mechanisms.

Melisma: the stretching of a single written syllable over several musical notes.

Metrical rests: the introduction of silences to draw breath or follow a drumbeat, producing acoustic gaps that have no equivalent in the manuscript.

Isochrony: the forcing of syllables of differing natural length into equal musical beats, overriding the prosody of the spoken language.

Accentual displacement and suppression: the movement or erasure of lexical stresses to accommodate a melody. In the data presented by Clarissa Vierke, which relies on Ernst Dammann's transcription of stanza DA 159, *kwa mámbo bora adhímu* (Vierke, 2011, p. 105), the execution transcribed, most likely chanted, leaves the penultimate stress of the disyllabic *bora* unrealized. The result is an execution with exactly two accented beats, *mámbo … adhímu*, which departs from the accentual structure the words supply. As section 3.3.3 establishes, such departures are not isolated: one invariant line can be projected into several contradictory accentual realities depending on the melody (*mahadhi*) chosen by the singer.

The formal reading is retained because it limits this variation, not because it removes it. The metrical grid of the composer is studied on the page. The reading serves another purpose: to observe what a speaker, instructed not to sing, realizes of the lexical stress structure, against which the accentual hypotheses are tested.

9.3.1. The Acoustic Method Illustrated: *Kiswahili!!*

The illustration uses the *shairi* poem *Kiswahili!!*, composed by Sheikh Mahmud Ahmed Abdull Kadir, Ustadh Mau (Abdulkadir & Frankl, 2013).

A chronological objection must be met first. *Kiswahili!!* is a twenty-first-century composition, and it is used here as a methodological illustration rather than as evidence about the classical corpus. Its interest lies in the rarity of a high-fidelity authorial recording, which allows the acoustic procedure to be shown on a text whose reader is also its composer. The case studies dedicated to the *shairi* in section 9.6 work on nineteenth-century compositions.

Two properties make the text suitable for this demonstration.

Philological stability. The written text has been published in both Ajami and Latin script (Abdulkadir & Frankl, 2013), and the version read by the poet in Roy (2013b, pp. 48–53), reproduced in Appendix A.2; the differences between the two, and between each and the reading, are listed in section 9.6.2. The recording was made in situ in Lamu in September–October 2005, during our fieldwork, and the phonetic description drawn from it was published in *KIAMU, archipel de Lamu (Kenya)* (Roy, 2013a).

Acoustic authenticity. The recording is of the *mtunzi* reciting his own composition. He reads and does not sing, so that Layer 3 is set aside and the signal remains that of the spoken language.

1. How the segmentation is established

The procedure has two steps, and they are not interchangeable.

The syllable boundaries are established by the researcher, in the Praat interface, by ear and by eye. The audio is looped, the auditory perception is cross-referenced with the visual data — specifically the onsets and peaks of the intensity contour — and the starting point of each syllabic nucleus is located, whether that nucleus is a vowel or a sonorous syllabic nasal. Nothing is automated at this stage, and nothing is inferred: the boundary is a decision, taken by a listener looking at a curve.

The curves are then plotted from those boundaries, in Python with Parselmouth. Once the timings are fixed they are written into the script, which recomputes the acoustic parameters from the original audio and renders the plot. Parselmouth interfaces directly with Praat's acoustic engine, so the measurement is the same; what the script adds is a standardized output, tuned pitch constraints that prevent octave-jump artifacts around voiceless consonants, and a figure that can be regenerated by any reader from the deposited audio.

A sixteen-syllable line compresses badly in a single frame, so each *shairi* line is plotted as its two hemistiches (*vipande*), which raises the temporal resolution and allows the syllabic grid to be aligned precisely.

2. The line examined

Stanza 3, line 1, in the numbering of Appendix A.2 :

Mimi mamenu sitasa, wala sina punguwani.

The line divides as 8 + 8.

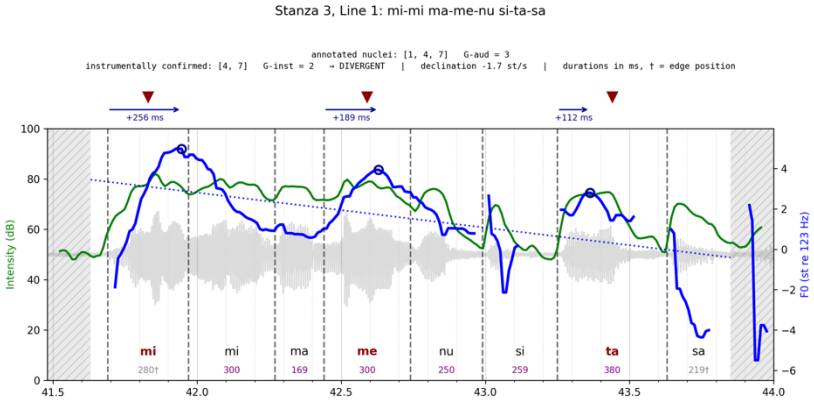
One feature of the first hemistich deserves a word before the measurement. *Mamenu* is a contraction of *mama yenu*, four syllables reduced to three. The composer adjusts the morphology so that the hemistich closes on exactly eight syllables; the count is not adjusted to the words.

2.a. Key to the figures

Green: intensity. Blue: F0 in semitones relative to the speaker median, measured once over the whole recitation so that all figures of a recitation share one scale. Dotted blue: the declination internal to the line, removed before measurement. Red triangles mark the

accentual nuclei established by ear; a question mark above a triangle marks a nucleus heard but judged weak. Open circles mark the measured F0 peak of each annotated nucleus, with its latency from the syllable onset. Durations are in milliseconds; a dagger marks the first and last syllables, excluded from durational comparison. Each caption gives the two values specific to its figure: the speaker median and the declination slope.

2.b. Figure 1 — first hemistich

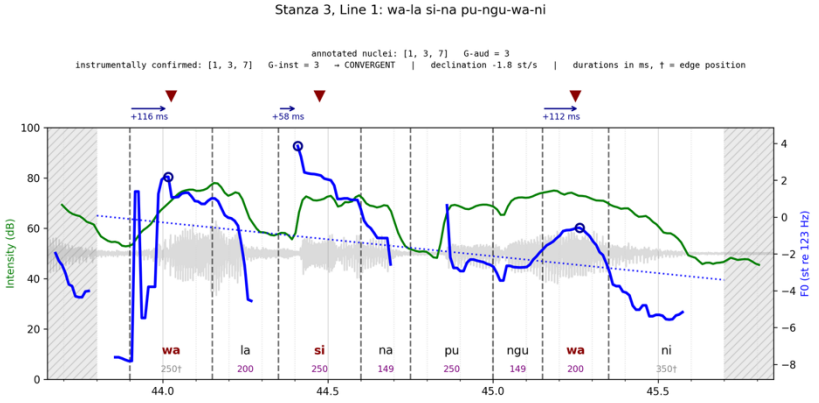


Kiswahili!!, Stanza 3, Line 1, H1: "mi-mi ma-me-nu si-ta-sa"
Lexical reading: 2·3·3, P_1=1, P_2=4, P_3=7, G-lex = 3.
Auditory reading: 2·3·3, P_1=1, P_2=4, P_3=7, G-aud = 3, delta —.
Instrumental reading: 5·3, P_1=4, P_2=7, G-inst = 2, delta $-P_1$=1 (against the lexical reading).
Speaker median 123 Hz; declination –1.74 st/s.

The instrumental reading diverges from the auditory reading: position(s) [1] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 mi      280 ms†  F0 -0.22 st  peak +256 ms  intensity  74.6 dB
  2 mi      300 ms   F0 -0.08 st  peak + 16 ms  intensity  76.5 dB
  3 ma      169 ms   F0 -1.35 st  peak +168 ms  intensity  75.0 dB
  4 me      300 ms   F0 +1.15 st  peak +189 ms  intensity  75.3 dB
  5 nu      250 ms   F0 +0.18 st  peak +  0 ms  intensity  65.7 dB
  6 si      259 ms   F0 -1.07 st  peak + 21 ms  intensity  56.0 dB
  7 ta      380 ms   F0 +1.61 st  peak +112 ms  intensity  65.7 dB
  8 sa      219 ms†  F0 -2.59 st  peak + 14 ms  intensity  63.1 dB
```

2.c. Figure 2 — second hemistich



Kiswahili!!, Stanza 3, Line 1, H2: "wa-la si-na pu-ngu-wa-ni"
Lexical reading: 2·2·4, P_1=1, P_2=3, P_3=7, G-lex = 3.
Auditory reading: 2·2·4, P_1=1, P_2=3, P_3=7, G-aud = 3, delta —.
Instrumental reading: 2·2·4, P_1=1, P_2=3, P_3=7, G-inst = 3, delta — (against the lexical reading).
Speaker median 123 Hz; declination –1.84 st/s.

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 wa     250 ms†  F0 +0.08 st  peak +116 ms  intensity  68.6 dB
  2 la     200 ms   F0 -0.27 st  peak +  7 ms  intensity  67.0 dB
  3 si     250 ms   F0 +3.03 st  peak + 58 ms  intensity  68.1 dB
  4 na     149 ms   F0 +0.46 st  peak +  9 ms  intensity  64.4 dB
  5 pu     250 ms   F0 -0.32 st  peak +110 ms  intensity  61.3 dB
  6 ngu    149 ms   F0 -0.44 st  peak +141 ms  intensity  70.5 dB
  7 wa     200 ms   F0 +1.38 st  peak +112 ms  intensity  73.0 dB
  8 ni     350 ms†  F0 -1.81 st  peak +  2 ms  intensity  57.4 dB
```

2.d. What the two hemistiches show

The second hemistich, *wala sina punguwani*. The lexicon yields 2·2·4 with stresses on positions 1, 3 and 7. The ear realizes that structure exactly. The instrument confirms it: the three prominences are the three highest detrended F0 means of the segment, at +0.08, +3.03 and +1.38 semitones. Three procedures, applied independently, return the same accentual structure, and that structure has three groups.

The first hemistich, *mimi mamenu sitasa*. The lexicon yields 2·3·3 with stresses on positions 1, 4 and 7, and the ear realizes that structure exactly. The instrument confirms only two of the three: it drops the nucleus on position 1, returning 5·3 on positions 4 and 7.

The margin on which that divergence rests. The syllable the ear hears as prominent, the first *mi*, measures −0.22 semitones; the syllable that follows measures −0.08. The difference is **0.14 semitones**, and it decides whether the first accentual nucleus of the hemistich exists. In the second hemistich the corresponding margin is 0.35 semitones, in the other direction. Two hemistiches of one line, delivered in a single breath by one speaker and measured by one procedure, receive opposite verdicts on a difference smaller than a third of a semitone.

The mechanism is identifiable. The F0 peak associated with the first *mi* arrives **256 milliseconds after the syllable onset**, the longest latency recorded anywhere in this volume. The melodic gesture initiated by the stressed syllable culminates well inside the syllable that follows it. The ear attributes the movement to the syllable that triggers it; a contour read syllable by syllable attributes it to the syllable in which it peaks. Neither procedure is mistaken about the signal; they resolve the same physical event differently.

3. What this line tests

The epistemological scope must be stated. One line of a twentieth-century corpus is not evidence on a macro-scale. It shows how the acoustic data is constructed and measured, and how it meets the conditions of section 9.2.4. Because the meter is *shairi*, the falsification bears on the framework formalized in section 2.5.2.

Conditions 1 and 3. The intensity contour delineates eight syllabic nuclei in the first hemistich and eight in the second. The line fulfils the 8 + 8 matrix without any rest or performative adjustment, satisfying Condition 1. No word crosses the boundary between the hemistiches, satisfying Condition 3.

Condition 5. The second hemistich carries three accentual groups on all three readings — lexical, auditory and instrumental. An axiom admits no exception, and this hemistich is one: the eight-syllable segment is not governed by a bipartite engine here, and the result does not depend on which procedure is trusted, since all three agree. Minerba's framework transposes Vierke's hypothesis, formulated for the eight-syllable *utenzi* line, onto the 2 × 8 structure of the *shairi* without structural validation.

Condition 6. The accentual peaks fall on positions 1, 4 and 7 in the first hemistich and 1, 3 and 7 in the second. In neither does the line exhibit the $P_1 = 3$, $P_2 = 7$ configuration, and the prototype is not what governs it.

And what the divergence between the two hemistiches establishes. The instrument confirms three groups in one hemistich and two in the other, on segments that the lexicon and the ear both read as having three. Three consequences follow, and they are not of equal weight.

The first is that the bipartite axiom is contradicted by a hemistich on which no reading dissents.

The second is that an instrumental reading can contradict the lexicon and the ear together, and does so here on a margin of 0.14 semitones. A procedure whose verdict turns on such a margin is not a procedure from which a metrical rule can be read off.

The third is the general one, and it is the reason this section opens the chapter rather than closing it. A meter is not derived from a signal. The signal records the execution of a text by a speaker at one moment, with the peak delays, declinations and boundary effects that physiology imposes; the meter is the constraint the composer worked under, and it is legible in the syllable count and the rhyme boundaries,

which the signal corroborates but does not constitute. The accentual tradition proceeds in the opposite direction, reading the beats off an execution and calling the result a matrix. The two hemistiches of this single line show what that direction costs: the same structure, measured the same way, yields two different answers.

Why a single line settles nothing else. This readout contradicts the bipartite axiom and shows no *Grenzschema*, but it rests on sixteen syllables. Localized acoustic extraction has to be coupled with the exhaustive mapping and the statistical tests of the case studies, which establish whether these configurations recur or fluctuate from line to line. That is the work of sections 9.4 to 9.6.

4. Reproduction

The figures of this line, like every figure of this kind in the volume, are produced by a script that recomputes them from the recording, here the poet's reading of *Kiswahili!!*, deposited in open access on Zenodo (Roy & ibn Kadir, 2016). Each script carries in its header the accentual analysis of its segment, the boundaries of the segment in the recording, and the onset of every syllable, and it regenerates its figure and its caption from the audio alone. The scripts are published on GitHub and archived on Zenodo, together with a Docker image that allows them to be run over the long term; these deposits, and the recordings of the classical corpus, are listed with their addresses and DOIs in section 9.2.7.

9.4. Utenzi Texts

9.4.1. Case Study I: *Utenzi wa Rasi 'lGhuli* — Mgeni bin Faqihi (c. 1850)

1. Epistemological Context and Experimental Protocol

To execute a definitive empirical robustness test of our formal models, we must select a dataset that is both monumentally classic and anchored in the oral, performative reality of the Swahili coast. The *Utenzi wa Rasi 'lGhuli* (*The Epic of the Ghoul's Head*), composed around 1850 in Bagamoyo by Mwalimu Mgeni bin Faqihi, perfectly fulfills these conditions (Mgeni bin Faqihi, 1979). As explicitly documented in the editorial prefaces of the manuscript, this text—

with its 4,584 stanzas (*beti*)—is the second-longest known Swahili epic. While the narrative is inspired by the Arabic فتوح اليمن *Futuhu 'lYamani* (recounting the conquest of Yemen by Muslim armies), the epic itself is not a mere translation. It is a sovereign, endogenous Swahili creation (Mgeni bin Faqihi, 1979, p. v), specifically crafted because the Arabic source was incomprehensible to the general coastal population.

Crucially for our falsification protocol against the "performative fallacy," the historical record confirms that this text was actively and regularly sung or read (*hughaniwa au kusomwa*) at communal events such as funerals (*matanga*) and weddings (*harusi*). The text documents famous reciters (*waghani*) like Mwalimu Sefu bin Himidi of Bagamoyo, who was invited as far as Tabora and Ujiji to sing this epic, alongside other masters like Ramdhani Songoro and Mwalimu Kisira Bakari (Mgeni bin Faqihi, 1979, p. [vi]). This indicates that the text we are analyzing is the exact architectural blueprint utilized by classical performers in the living tradition.

To resolve the metrical debate and avoid mistaking the singer's musical melody *(mahadhi)* for the composer's structural metrical grid *(arudhi)*, our methodology relies on lexical, auditory, and acoustic evidence, according to the three-step protocol defined in Section 9.2.3. The dataset for this experiment comprises a strictly controlled block of the first 40 stanzas *(beti)* of the epic—yielding 160 graphic lines.

To guarantee the empirical neutrality of this test, we collected our acoustic data by recording a formal, spoken recitation (rather than a sung performance) of the manuscript by our L1 Standard Swahili speaker. The resulting audio track was subsequently processed using Praat, the industry-standard software for acoustic phonetics. Following the tripartite protocol established in Section 9.2.3, the precise segmentation of syllabic boundaries for the instrumental sample was conducted through rigorous manual verification on the waveforms, whereas the accentual prominences were established by rule and by ear prior to any instrumental measurement.

Syllabic Mapping *(Vina na Mizani)*: For the instrumental sample, we explicitly mapped the exact syllabic divisions by combining critical auditory listening with direct visual identification of the syllable boundaries on the Praat waveforms. A Swahili syllable corresponds to a measurable acoustic event: we auditorily verified and visually tracked the waveform oscillations (the gray oscillogram in the figures) and the characteristic intensity peaks to scientifically confirm

our auditory perception through the visual, physical presence of exactly 8 syllables per line in the acoustic signal.

Accentual Mapping: The accentual structure of the line is deterministically derived from the lexicon and exhaustively verified by ear across the entire corpus. The instrumental measurement is applied only to a sample. The instrumental reading algorithmically designates acoustic prominence based strictly on the contour of the Fundamental Frequency (F_0), represented by the blue curve in the figures, following standard practices in Swahili linguistics. While relative intensity amplitude (the green curve) and vocalic duration (measured along the horizontal axis) are also extracted and systematically reported for each syllable, they are not decisive on their own and are not used by our algorithm to designate the accentual peaks.

In accordance with the symbolic alphabet defined in Section 2.5.1, the accentual structure is represented by a sequence of syllabic volumes comprising each accentual nucleus (e.g., $2 \cdot 2 \cdot 2 \cdot 2$ or $2 \cdot 6$), and the positional index of each stressed syllable within the 8-syllable line is noted as P_i. The total number of accentual groups is recorded for each of the three readings as a distinct parameter—yielding G-lex for the lexical baseline, G-aud for the auditory realization, and G-inst for the instrumental extraction.

For the instrumental verification, any divergence between the machine's F_0 extraction and the established readings is recorded as a physical fact, without revising the initial annotation. Specifically, the formal delta (e.g., $-P_n$ for an expected prominence not designated by the instrument, or $+P_n$ for a measured prominence not predicted by the lexicon) is calculated mechanically against the deterministic lexical baseline (G-lex). However, the acoustic figures visually plot the auditory annotation (G-aud, marked by red triangles) against the instrumental measurement (G-inst, marked by open circles on the F_0 peak) to explicitly declare whether the physical signal confirms or diverges from the human ear.

2. Methodological Caveat: The Arabic Definite Article ('l)

Before proceeding to the formal analysis, a strict methodological filter must be applied regarding the Arabic definite article (the *'l* particle, representing the Arabic *al-*). In the editorial notes detailing the transcription symbols of the manuscript, it is explicitly stated that

this particle: "indicates the letter of definition which is the Arabic 'article' [...] It is always a distinct syllable" (Mgeni bin Faqihi, 1979, p. xxiv)[85].

However, the empirical auditory analysis of the recitation reveals a contradiction with this editorial claim, demonstrating a phenomenon of syllabic elision during the natural flow of reading. For instance, in Stanza 1, Line 2 (ji-na la Mo-la 'l-la-hi), the particle is indeed pronounced as a distinct syllable (sounding like /il/) and counts perfectly as 1 syllable to reach the 8-syllable boundary. Yet, in Line 1 of the exact same stanza (A-wa-li bi-su-mi 'lla-hi), treating the particle as a distinct syllable would yield 9 syllables. In practice, the reader automatically elides it, reducing its syllabic value to 0. The resulting auditory realization naturally maintains the 8-syllable *mizani*.

While these elisions may demonstrate the absolute supremacy of the 8-syllable grid—forcing Arabic loanwords to compress or expand to fit the Bantu matrix—they introduce a localized phonological ambiguity. To maintain absolute scientific rigor and to prevent accentualist theorists from exploiting these specific Arabic morphological remnants to contest our accentual mapping, any line containing the *'l* particle is strictly excluded from our statistical falsification pool. We will note and syllabify them in the sample to observe the grid (marking them [Excl.]), but our definitive Popperian test will rest exclusively on the unambiguous Swahili lines.

Additionally, in the reading text provided to our reciter, Fatuma Shaaban, the original letter *q* (representing the Arabic *qaf*, ق) was transcribed as *k*. This practical adjustment was made to avoid disrupting her natural reading flow, as the letter *q* is absent from Standard Swahili orthography.

Despite this orthographic adaptation, Fatuma Shaaban, as an L2 speaker of Arabic, naturally articulates certain Arabic phonemes during her reading. When she encounters a highly recognizable Arabic loanword, she applies the original Arabic realization to specific consonants — the voiceless pharyngeal fricative (*ḥā'*, ح) in *Habibu*, for instance.

A few of these realizations are documented below, and the reason for noting them must be stated exactly. They are provided for transparency, to clarify acoustic features that a reader might otherwise mistake for divergences from the source text. In the specific

[85] Swahili original: "yaonyesha harfu 'ltaarifu iliyo "article" ya Kiarabu [...] Ni daima silabi ya pekee".

instances recorded here, it is simply observed that their occurrence does not alter the formal syllable count *(mizani)*. This list is purely indicative and does not constitute an inventory of the Arabic phonology of this reading. A systematic survey of Arabic phonological realizations falls outside the scope of this metrical analysis, and no exhaustiveness is asserted. The raw audio recordings are deposited in full, so that a reader qualified to conduct such a survey may do so.

3. The Competing Formal Models

Before the corpus is confronted with either model, we define the exact form of the two competing theories to be tested against the lexical and auditory structures established above.

3.a. The Endogenous Syllabic Matrix

As formalized in Chapter 8, the *utenzi* stanza is governed exclusively by the absolute syllable count (*mizani*) and the rhyme boundaries (*vina*). The stanza is a block of four lines, each of eight syllables. The first three share a local rhyme *a*, whose scope is limited to the stanza and which changes from one stanza to the next; the fourth carries the global rhyme *X*, the *bahari*, which spans the whole poem. Both variables are written unresolved in the matrix, and each is governed by the hierarchy defined in the symbolic alphabet, $a^s > a^v$ and $X^s > X^v$: the syllabic form is the rule and is tested first, the vocalic form a license examined only where the syllabic one has been contradicted.

Formal Parameters: $n_{max} = 4, H = 1, S = 8$

The Formal Matrix:
L1: [8, a] //
L2: [8, a] //
L3: [8, a] //
L4: [8, X] //

The Algebraic Equation:
$\boldsymbol{Utenzi} = \mathbf{3}(\boldsymbol{N_a}//) \cdot \boldsymbol{N_X}//$ (where N=8)

3.b. The Exogenous Accentual Model / Vierke's *Grenzschema*

As examined in Chapter 2, the accentual school posits that the 8-syllable line is inherently bipartite, driven by an exogenous accentual superstructure called the *Grenzschema* (see Sections 2.3.1 and 2.5.1).

The Axiom of Two Beats: $G = 2$ (every line strictly contains exactly two metrical beats, hence two accentual groups).

The *Grenzschema*: The cognitive prototype is a perfectly symmetrical line where the two stresses are locked on specific positional indices: $P_1 = 3$ and $P_2 = 7$.

The Accentual Matrix ($Matrix_{ideal}$):
L1: [u u - u] ·[u u - u] //
L2: [u u - u] ·[u u - u] //
L3: [u u - u] ·[u u - u] //
L4: [u u - u] ·[u u - u] //

Formal Sequence Translation: $4 \cdot 4$, $P_1 = 3$, $P_2 = 7$, $G = 2$.

As established above, this exogenous theoretical model is tested exclusively against our exhaustive lexical (G-lex) and auditory (G-aud)readings, while the instrumental reading (G-inst) serves only as a localized, purely indicative physical verification on the sample.

As detailed in Section 2.3, the *utenzi* meter, as theorized by Vierke, operates as a mixed meter (*gemischt-quantifizierend*). It fully adopts the baseline syllabic framework discussed above, grafting an accentual metric onto it.

4. Utenzi wa Rasi 'lGhuli: Exhaustive Syllabic and Accentual Mapping, Stanzas 1–40

4.a. Reading the entries

Each line occupies a block of four short lines, in the notation of sections QRG.1 and 2.5.

The first gives the reference and the line as it stands in the source text supplied to the reader, syllabified, with its syllable count; *guni* marks a line whose count departs from eight in the source itself. The second is the lexical reading, derived from that text by the penultimate rule and requiring no recording: the sequence of nucleus volumes

|A1|·|A2|·…, the positional index P*i* of the stressed syllable of each nucleus A*i*, and G-lex.

The third line appears only where the execution departs from the source, and gives the line as the reader actually pronounced it, syllabified, with the count her execution yields. The fourth is the auditory reading, established by ear on that execution, followed by Δ, which records the transition from the lexical structure. The notation −P*i*=n marks the lexical nucleus A*i*, whose stress falls on syllable n, as unrealized; +P*i*=n marks a realized nucleus A*i*, with its stress on syllable n, that the lexicon does not predict; a dash marks a realization that follows the lexical structure exactly. Where the line is excluded from the statistical pool, this line gives the ground of exclusion instead.

A question mark on a positional index marks an auditory decision made without full confidence: P2=5? is a nucleus heard but judged weak and counted nonetheless. The position is counted as written, and the mark registers the doubt without resolving it. No such mark can appear in the lexical reading, which is derived by rule.

A nucleus of a single syllable occurs twice in the auditory reading and never in the lexical one. The penultimate rule cannot produce it, since a monosyllabic phonological word is proclitic and carries no lexical stress. What the reading records in these two lines is a prominence placed on a monosyllable by the speaker, which is the oratorical accent Sacleux described in 1909 as free in its placement and belonging to delivery rather than to the text. It is registered as an addition and counted as such.

4.b. Mapping

Stanza 1
1.1 a-wa-li bi-su-mi 'lla-hi (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
excluded — Arabized execution
1.2 ji-na la mo-la il-la-hi (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
excluded — Arabized execution
1.3 pwe-ke a-si-ye sha-bi-hi (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ −P2=4

1.4 ndi-ye wa-hi-di ka-ha-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 2
2.1 a-ra-h-ma-ni ra-hi-mi (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
execution a-rah-ma-ni ra-hi-mi (7)
excluded — Arabized execution
2.2 ndi-ye ha-i 'l-ka-yu-mi (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
excluded — Arabized execution
2.3 mwe-nye e-zi ya da-wa-mi (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —
2.4 a-li-pe-nda-lo hu-ji-ri (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
Stanza 3
3.1 hu-wa la-hu 'l-ja-la-li (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
execution wa-lla hu-ja-la-li (6)
excluded — Arabized execution
3.2 ndi-ye mu-i-zi mu-dhi-li (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
3.3 ndi-ye ha-ki-mu a-di-li (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
3.4 ndi-ye m-ju-vi wa si-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 4
4.1 hu-wa 'l-la-hu ma-na-ni (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
execution hu-wa-lla hu ma-na-ni (7)
excluded — Arabized execution
4.2 m-tu-ku-fu mwe-nye sha-ni (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —

4.3 kha-li-ku a-li-nsa-ni (7)
lexical 3·4 — P1=2, P2=6 — G-lex 2
execution kha-li-ku a-li-nsa-ni — false start in the pronunciation
excluded — false start
4.4 bi u-na-si wa dhu-ku-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 5
5.1 hu-wa 'l-la-hu wa-ba-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
execution hu-wa-lla hu-wa-ba-ri (7)
excluded — Arabized execution
5.2 wa-hi-di 'l-ka-ba-ri (7)
lexical 3·4 — P1=2, P2=6 — G-lex 2
execution wa-hi-di 'l-ka-ba-ri (7)
excluded — Arabized execution
5.3 'l-ka-bi-ri 'l-gha-fa-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
execution ul-ka-bi-ri ul-gha-fa-ri (8)
excluded — Arabized execution
5.4 mu-da-bi-ri li u-mu-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
execution mu-dab-bi-ril u-mu-ri (7)
excluded — Arabized execution
Stanza 6
6.1 ndi-ye ka-ri-bu m-ji-bu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
6.2 ndi-ye gha-fi-ri dhu-nu-bu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
6.3 ndi-ye sa-ta-ri a-i-bu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
execution ndi-ye sa-ta-ril a-i-bu (8) — Arabic realization of the *ain* in *aibu*; no change of count
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
6.4 ndi-ye ra-hi-mu shu-ku-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —

Stanza 7
7.1 hu-wa la-hu 'l-kha-li-ki (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
execution hu-wa-lla hul-kha-li-ki (7)
excluded — Arabized execution
7.2 ndi-ye mu-u-mba ri-zi-ki (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
7.3 ku-we-ne-za ma-khu-lu-ki (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
execution ku-e-ne-za ma-khu-lu-ki (8)
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
7.4 wa-ku-u ha-ta sa-ghi-ri (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
Stanza 8
8.1 sha-ra-ka wa gha-ri-bi-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
8.2 su-he-li wa ki-bi-li-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
8.3 ka-we-ne-za wo-te pi-a (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
execution ku-e-ne-za wo-te pi-a (8) — audible swallow after *kueneza*
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
8.4 wa-le ba-ra na ba-ha-ri (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
Stanza 9
9.1 al-la al-la wa-li-mwe-ngu (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
execution a-la a-la wa-li-mwe-ngu (8)
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —
9.2 tu-mu-hi-mi-di mu-u-ngu (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —

9.3 mu-u-mba i-nti na mbi-ngu (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
execution mu-u-mba n-ti na mbi-ngu (8)
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ –P2=4
9.4 ma-ja-ba-li na sha-ja-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
execution n-ja-ba-li na sha-ja-ri (8) — *Majabali* reduced to *njabali*; no metrical consequence
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 10
10.1 tu-mu-hi-mi-di ja-la-li (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
10.2 ni ja-wa-bu a-fa-dha-li (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
10.3 ku-m-sa-li-a ra-su-li (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
10.4 ndi-ye ja-wa-bu khi-a-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 11
11.1 tu-m-sa-li-e m-sha-ra-fu (9) — *guni*
lexical 5·4 — P1=4, P2=8 — G-lex 2
auditory 5·4 — P1=4, P2=8 — G-aud 2 — Δ —
11.2 tu-me wa mo-la la-ti-fu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
11.3 ke-sho tu-pa-te wo-ko-fu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
11.4 twe-pu-ka-ne na sa-i-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4.4 — P1=3, P2=7 — G-aud 2 — Δ —

Stanza 12
12.1 tu-m-sa-li-e ha-bi-bu (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
execution tu-m-sa-li-e ha-bi-bu — Arabic realization of *ḥā*ʾ in *Habibu*; no change of count
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
12.2 a-sa i-ka-wa na-si-bu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
12.3 kwe-pu-ka-na na a-dha-bu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
12.4 hi-yo si-ku ya nu-shu-ri (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2.4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —
Stanza 13
13.1 ni-ko-mi-le ku-ra-di-di (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
13.2 ni-m-shu-ku-ru wa-du-di (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
13.3 na si-fa za mu-ha-ma-di (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution na si-fa za mu-ha-ma-di — Arabic realization of *ḥā*ʾ in *Muhamadi*; no change of count
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
13.4 ku-m-si-fu si-ka-di-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 14
14.1 e-we mwa-na-ngu kha-mi-si (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
execution e-mwa-na-ngu kha-mi-si (7)
excluded — *guni* of reading
14.2 ni-pa-ti-e ka-ra-ta-si (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

14.3 u-ni-le-te-e u-pe-si (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
14.4 pa-si-we ku-ta-a-kha-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 15
15.1 ka-ra-ta-si i-we sha-mu (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
15.2 i-li-yo nje-ma na-i-mu (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
15.3 na pa-mbo ni ma-a-di-mu (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
15.4 twa-a hi-yo ya ha-ri-ri (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —
Stanza 16
16.1 we-nda-po u-ki-nu-nu-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
16.2 ji-ta-hi-di ku-te-u-a (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
16.3 na-we ni m-tu mwe-le-wa (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
16.4 si m-cha-che wa na-dha-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 1·3·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ +P1=1
Stanza 17
17.1 u-na-po-kwe-nda du-ka-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
17.2 u-si-u-si-ri ndi-a-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —

17.3 u-fi-ke kwa se-le-ma-ni (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
17.4 a-u kwa i-sa m-zu-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ −P1=1
Stanza 18
18.1 m-pa-ti-e na ma-pe-sa (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
18.2 u-ya-tu-ku-e ka-bi-sa (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
18.3 u-ki-wa-si-li kwa i-sa (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
18.4 u-si-we ku-ta-a-sa-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 19
19.1 cha-mba hu-yo se-le-ma-ni (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ −P1=1
19.2 hu-ku-mu-o-na du-ka-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
19.3 na u-ru-di i-yo hi-ni (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ −P2=5
19.4 e-nde kwa i-sa m-zu-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 20
20.1 du-ka-ni u-ta-po-ru-di (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —

20.2 na sa-fi-di-ye ma-da-di (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
execution na sa-di-fi-ye ma-da-di (8)
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
20.3 na ndu-gu-yo ma-su-u-di (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
20.4 o-go-pa ku-mu-da-bi-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 21
21.1 u-ta-po-ru-di du-ka-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
21.2 u-ki-wa-si-li nyu-mba-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ +P1=2
21.3 u-si-we ku-ta-ma-ka-ni (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
21.4 u-sa-fi-di zi-ngi-fu-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
execution u-sa-fi-di si-zi-fu-ri (8) — the reader transposes *zi* and *si*; no metrical consequence
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 22
22.1 kwa pa-po u-ya-te-u-e (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
22.2 za-yo-ba-ki u-ya-to-e (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
22.3 u-sa-fi ha-ta u-ju-e (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ −P2=4
22.4 i-ti-wa-po i-na-wi-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

Stanza 23
23.1 u-ni-pa-ti-e ki-ba-o (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
23.2 ki-zu-ri ki-si-o ta-o (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ –P2=5
23.3 na u-zi na u-je na-o (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
23.4 u-li-o mwe-ma khi-ya-ri (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ –P1=2
Stanza 24
24.1 na u-zi u-we na ta-si (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
24.2 u-po-te u-ngi-e ka-si (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
24.3 hi-pi-gwa i-we na-fa-si (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
execution i-pi-gwe i-we na-fa-si (8)
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ –P2=4
24.4 i-tu-li-ye mi-sta-ri (7) — *guni*
lexical 4·3 — P1=3, P2=6 — G-lex 2
auditory 4·3 — P1=3, P2=6 — G-aud 2 — Δ —
Stanza 25
25.1 pa-na ha-di-thi a-ja-bu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
25.2 ya-ndi-kwa nda-ni ki-ta-bu (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
25.3 kwa lu-gha ya wa-a-ra-bu (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —

25.4 na-yo ha-di-thi khi-ya-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 26
26.1 ha-di-thi hi-yo yu-a-ni (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
26.2 i-na ma-ne-no ba-ya-ni (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
26.3 ya vi-ta vya al-ya-ma-ni (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution ya vi-ta vya i-ya-ma-ni (8)
excluded — 'l particle
26.4 za-ma-ni za-ke ba-shi-ri (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 27
27.1 ha-i-o-na ha-i-so-ma (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
27.2 ma-a-na ni ku-fa-ha-ma (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
27.3 ha-ti-bu wa-ngu m-ti-ma (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ −P2=4
27.4 hu-wa-za na ku-fi-ki-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 28
28.1 hu-fi-ki-ri ni-ki-we-za (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
28.2 ha-o-na ku-ni-pe-nde-za (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution na-ha-o-na ku-ni-pe-nde-za (9)
excluded — *guni* of reading

28.3 mo-yo-ni ni-ki-le-ke-za (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
28.4 ku-tu-nga kwa u-sha-i-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 29
29.1 ha-pe-nda ku-i-ba-di-li (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
29.2 kwa lu-gha ya su-wa-hi-li (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
29.3 ki-a-ra-bu ni ka-li-li (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
29.4 wa-ju-a-o taf-si-ri (7) — *guni*
lexical 4·3 — P1=3, P2=6 — G-lex 2
execution wa-ju-a-o taf-si-ri (7) — the reader executes *taf-si-ri*, retaining the Arabic closed syllable; the shortened form is that of the source
auditory 4·3 — P1=3, P2=6 — G-aud 2 — Δ —
Stanza 30
30.1 ni-ka-wa ku-ji-ka-li-fu (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
30.2 ki-a-ra-bu ha-wa-ki-fu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
30.3 ma-ne-no ha-ya-sa-ni-fu (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 2
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
30.4 ku-wa-pa zo-te ha-ba-ri (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 31
31.1 na-che-le-a ni m-ji-nga (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

31.2 si-ju-i sa-na ku-tu-nga (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
31.3 na ma-ne-no ku-ya-pa-nga (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
31.4 ya-ta-ka m-tu ma-hi-ri (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
Stanza 32
32.1 wa-la mi si mu-a-li-mu (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 3·1·4 — P1=2, P2=4, P3=7 — G-aud 3 — Δ −P1=1, +P1=2, +P2=4
32.2 wa-la si-ju-i i-li-mu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
32.3 na-ta-ka ku-ta-ya ma-mu (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
32.4 zi-di-ni ku-ni-u-dhu-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 33
33.1 m-si-je ku-o-na ki-na (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2?, P2=5, P3=7 — G-aud 3 — Δ —
33.2 ki-si-cho-shi-ka ma-a-na (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
execution ki-si-sho-shi-ka ma-a-na (8) — whispered sibilant in *kisichoshika*
auditory 5.3 — P1=4, P2=7 — G-aud 2 — Δ —
33.3 m-ki-wa-mo ku-ni-ne-na (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
33.4 wa-ku-u ha-ta sa-ghi-ri (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —

Stanza 34
34.1 m-ka-na-ndi-ka a-i-bu (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
34.2 wa-la pa-si-po sa-ba-bu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
34.3 ku-ni-ru-di ni tha-wa-bu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
34.4 wa-la na-mi si-ka-si-ri (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
Stanza 35
35.1 e-we mwa-na-ngu kha-mi-si (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
execution e mwa-na-ngu kha-mi-si (7)
excluded — *guni* of reading
35.2 shi-ka hi-ma ka-ra-ta-si (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
35.3 na ka-la-mu kwa u-pe-si (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
35.4 pa-si-we na ku-u-si-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 36
36.1 o-go-pe ku-ta-ka-ra-bu (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
36.2 ni-kwa-mbi-e ma-tu-lu-bu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
36.3 ni-ku-pa-lo u-ku-tu-bu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

36.4 u-si-we ku-ta-a-kha-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution u-si-we ku-ta-a-kha-ri — Arabic realization of *kh* in *kutaakhari*; no change of count
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 37
37.1 hi-ma a-ndi-ka u-pe-si (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
37.2 su-tu-ri zi-we na-fa-si (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ –P2=4
37.3 na ma-ne-no u-ka-ki-si (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
37.4 u-sa-ndi-ke m-sha-dha-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 38
38.1 na kha-ti i-we fa-si-ha (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
execution na kha-ti i-we fa-si-ha — Arabic realization of *kh* and *ḥāʾ* in *khati* and *fasiha*; no change of count
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
38.2 ma-ne-no ya-we wa-dhi-ha (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
38.3 bi u-dhu-ma na fa-ta-ha (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
38.4 na ja-zi-ma na ka-si-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 39
39.1 i-wa-po u-zi-fa-ni-di (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution i-wa-po u-zi-fa-ni-di — false start on *uzifa*
excluded — false start

39.2 kwa ma-da na ta-shi-hi-di (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution kwa ma-da na ta-shi-di (7)
excluded — *guni* of reading
39.3 zi-we ba-i-di ba-i-di (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
execution zi-we ba-i-di ba-i-di — Arabic realization of the *ain* in *baidi*; no change of count
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
39.4 zi-si-we mu-da-a-tha-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ +P2=4
Stanza 40
40.1 na nu-ku-ta ti-a pi-a (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
40.2 u-si-wa-te ha-ta mo-ya (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
40.3 u-zi-pa-nge kwa sa-wi-a (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
40.4 zi-si-we ku-ta-gha-ya-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
execution zi-si-we ku-ta-gha-ya-ri — false start on *kutagha*
excluded — false start

4.c. Excluded lines

Nineteen lines are removed from the statistical pool, leaving 141. Arabized executions preserve closed syllables and geminates and thereby displace the stress positions mechanically; the article *'l* is the recurrent cause in the opening invocation. *Guni* of reading are lines in which the reader adds or omits a syllable that the source text does not contain. False starts are lines the reader began twice, and which cannot be verified. Genuine *guni* of the source text are retained, and there are three: 11.1 at nine syllables, 24.4 and 29.4 at seven, all three verified against the printed edition.

1.1 *a-wa-li bi-su-mi 'lla-hi* — Arabized execution
1.2 *ji-na la mo-la il-la-hi* — Arabized execution
2.1 *a-ra-h-ma-ni ra-hi-mi* — Arabized execution
2.2 *ndi-ye ha-i 'l-ka-yu-mi* — Arabized execution
3.1 *hu-wa la-hu 'l-ja-la-li* — Arabized execution
4.1 *hu-wa 'l-la-hu ma-na-ni* — Arabized execution
4.3 kha-li-ku a-li-nsa-ni — false start
5.1 *hu-wa 'l-la-hu wa-ba-ri* — Arabized execution
5.2 *wa-hi-di 'l-ka-ba-ri* — Arabized execution
5.3 *'l-ka-bi-ri 'l-gha-fa-ri* — Arabized execution
5.4 *mu-da-bi-ri li u-mu-ri* — Arabized execution
7.1 *hu-wa la-hu 'l-kha-li-ki* — Arabized execution
14.1 e-we mwa-na-ngu kha-mi-si — *guni* of reading
26.3 *ya vi-ta vya al-ya-ma-ni* — Arabized execution
28.2 *ha-o-na ku-ni-pe-nde-za* — *guni* of reading
35.1 e-we mwa-na-ngu kha-mi-si — *guni* of reading
39.1 i-wa-po u-zi-fa-ni-di — false start
39.2 kwa ma-da na ta-shi-hi-di — *guni* of reading
40.4 zi-si-we ku-ta-gha-ya-ri — false start

4.d. Summary

The pool comprises **141 lines** of the 160 mapped. The syllabic constant N = 8 holds in 138 of them, that is 97.9 percent; the three remaining lines are *guni* of the source text, retained in the pool.

Distribution of G-lex, derived from the printed lines: G = 2 in 79 lines (56.0 percent), G = 3 in 62 (44.0). No line yields one accentual group, and none yields four.

Distribution of G-aud, established by ear in the formal reading: G = 2 in 86 lines (61.0 percent), G = 3 in 55 (39.0).

The transitions comprise 12 unrealized nuclei and 5 added prominences, affecting 15 lines, that is 10.6 percent of the pool. The unrealized nuclei fall on position 4 (six occurrences), position 1 (three), position 5 (two) and position 2 (one). No nucleus on position 7 is ever unrealized. Nineteen lines carry a positional index marked as uncertain. The unrealized nuclei fall on position 4 (eight occurrences), position 1 (three), positions 3 and 5 (two each) and position 2 (one). **No nucleus on position 7 is ever unrealized.** Eighteen lines carry a positional index marked as uncertain.

The *Grenzschema* appears as a lexical sequence 4·4 in **32 lines (22.7 percent)**, and is realized as P1=3, P2=7 with G-aud = 2 in 33 lines **(23.4 percent)**.

Inter-pass control

The auditory annotation is the work of a single listener, and the corpus was annotated twice. The two passes are not, however, comparable in the way a repeated measurement is. The first was carried out years earlier, under a methodology and a notation that preceded those of the present study: it recorded a footprint in long and short marks, without the sequence of nucleus volumes, without the transition Δ, and without any means of registering an uncertain decision. The second was carried out under the notation set out in section 9.2.3, and every line was reannotated from the recording.

What can be compared is therefore the coincidence of two protocols and not the reproducibility of one. Of the 141 lines in the pool, the two annotations assign the same accentual positions in **119 lines, that is 84.4 percent**, and differ in 22. Among the divergent lines, the second annotation records fewer accentual groups than the first in 11 lines and more in 4, while in 7 the number is unchanged and only the positions differ. Nineteen lines carry a positional index marked as uncertain in the second annotation, a category that had no equivalent in the first.

The figure should be read for what it is. It measures the agreement between an annotation made under one notation and an annotation made under another, on the same recording and by the same listener. Part of the movement is attributable to the instrument: a notation that records hesitation rather than resolving it makes it easier both to retain a weak nucleus and to set one aside, and the eighteen uncertain indices are the visible trace of that. The direction of the divergence is nonetheless worth noting, since it runs against the interest of the argument advanced here: the second pass finds fewer accentual groups than the first three times more often than it finds more, which is to say it is the more favorable of the two to the bipartite axiom under test.

None of this substitutes for a second annotator, and no claim of inter-annotator reliability is made. The raw audio is deposited so that any reader may supply one.

5. Statistical Analysis and Falsification of the Accentual Model

All figures are computed on the pool of 141 lines established in section 4. G-lex is derived from the printed lines by the penultimate rule and requires no recording; G-aud is the realization established by ear. The prose baseline invoked below is that of section 9.7.

5.a. Executing Conditions 1 and 2: the syllabic matrix and the rhyme

Condition 1, the syllabic constant. The matrix requires S = 8 for every line. The constant holds in **138 of the 141 retained lines, that is 97.9 percent**. The three exceptions are *guni* of the source text: 11.1 at nine syllables, 24.4 and 29.4 at seven, all three verified against the printed edition. As the condition provides, they are deviations of the text and not of the execution, they are retained in the pool, and they count against the matrix.

The condition is satisfied on its stated criterion. Deviation is rare, and the deviant lines are *guni* in the endogenous sense. Nineteen further lines were removed from the pool for reasons of execution and not of composition: eleven Arabized executions concentrated in the opening invocation, where the Arabic article *'l* is restored with its own phonology; four *guni* of reading; three false starts; and one line whose execution of the article rendered it unusable. Every determination is made by direct comparison with the text supplied to the reader, published in full in section 4.

Condition 2, the rhyme variables. The matrix of section 3 leaves both rhyme variables unresolved, written *a* and X, and each is tested in the order its hierarchy prescribes.

The global constant is governed by $X^{s} > X^{v}$, so X^{s} is tested first. X^{s}, the syllabic global rhyme, requires that every anchored line of the poem close on the same full consonant-vowel sequence. It holds, in **40 stanzas of 40**: every fourth line of the mapped corpus closes on the syllable *-ri*, without a single exception. The stronger form is satisfied, no second test is required, and the global constant of this corpus is X^{s}.

The local variable is governed by $a^{s} > a^{v}$, and a^{s} is tested first. a^{s}, the syllabic local rhyme, requires that the first three lines of each stanza close on the same full syllable. It holds in **38 of the 40 stanzas**. Two stanzas contradict it: Stanza 16, where the three lines close on *ukinunua*, *kuteua* and *mwelewa*, and Stanza 40, where they close on

pia, *moya* and *sawia*. In both the terminal vowel is constant and the onset varies. a^{v} is therefore tested for those two stanzas, and it holds. The local rhyme of this poem is syllabic in thirty-eight stanzas and vocalic in two.

The proportion is itself a result and is reported as one. The syllabic form is the rule, and the composer meets it in nineteen stanzas out of twenty; the vocalic form is a license, taken where the lexicon offers no syllabically rhyming word of the required meaning and measure. Two stanzas out of forty measure the lexical pressure this composer met, and no more than that.

The matrix of this corpus therefore reads:

***Utenzi* = 3(8 a^{s} //) · 8 X^{s} //,** with a^{v} at Stanzas 16 and 40.

Conditions 1 and 2 are therefore satisfied, with three *guni* declared, a syllabic global rhyme holding without exception, and a syllabic local rhyme holding in thirty-eight stanzas of forty. It should be stated plainly what this establishes and what it does not. These conditions verify the exhaustiveness of a description derived from the endogenous theorists and applied to a text they did not analyze; they do not falsify a rival hypothesis. The falsification is carried by Conditions 5 to 8, which follow.

5.b. Executing Condition 5: the bipartite accentual axiom (G = 2)

Condition 5 asks whether the eight-syllable line carries exactly two accentual groups, as Axiom 1 of section 2.5.1 requires. The question is put twice, to the lexicon and to the reading.

The lexicon of the line yields exactly two accentual groups in **79 lines of 141, that is 56.0 percent**, and three in 62, that is 44.0 percent. No line yields one accentual group and none yields four. The formal reading realizes exactly two prominences in **86 lines, that is 61.0 percent**, and three in 55.

These figures must be stated without embellishment. The bipartite count is the outcome in more than half the lines on the lexical measure and in three lines out of five on the auditory one, and any account of this corpus which denied that two accents are frequent would be contradicted by its own data. No such claim is made here.

What the figures establish is that the count is not a law. As an axiom, G = 2 admits no exception, and this corpus supplies 62 counter-examples on the lexical measure and 55 on the auditory one. As a statistical tendency, it holds in three lines out of five and fails in two.

The transition disposes of the objection that the count is an artifact of the analyst's ear. Of the 79 lines whose lexicon already yields two groups, 75 are realized with two: where the text offers the bipartite configuration, the reading takes it. Of the 62 lines whose lexicon yields three, **only 11 are brought to two in the reading**, that is 17.7 percent; 51 are not. The voice does not drive the line toward a binary pulse; it very largely preserves the accentual structure the lexicon supplies.

The prose baseline settles what this movement means. In the control corpus of section 9.7, read by the same speaker under the same instruction, the reading departs from the lexical structure in 12.5 percent of lines; in this corpus it departs in 10.6 percent. The two figures are of the same order. A reader alters the accentual structure of roughly one line in nine, at a rate that does not distinguish a nineteenth-century *utenzi* from a sentence about the price of fish. The reduction of G-lex to G-aud is therefore not a poetic phenomenon and cannot be what a meter does to a line.

Deletions and additions. The transitions comprise 12 unrealized nuclei and 5 added prominences. The unrealized nuclei fall on position 4 (six occurrences), position 1 (three), position 5 (two) and position 2 (one). **No nucleus on position 7 is ever unrealized.** Among the additions, two are of a kind this corpus alone produces: at 16.4 and at 32.1 the reader places a prominence on a monosyllabic word, which the penultimate rule cannot assign, since a monosyllable is proclitic and carries no lexical stress. What is recorded there is the oratorical accent Sacleux described in 1909 as free in placement and belonging to delivery rather than to the text, and it is counted as an addition. 19 lines carry a positional index marked as uncertain.

5.c. Executing Condition 6: the *Grenzschema* as a cognitive prototype

Condition 6 concerns not the number of the groups but their volumes and positions. The *Grenzschema* is the ideal matrix of section 2.5.1: $|A_1| = 4$, $|A_2| = 4$, with $P_1 = 3$ and $P_2 = 7$, which is the single sequence 4·4.

The lexicon of this poem produces that sequence in **32 lines of 141, that is 22.7 percent**, and it is realized as $P_1 = 3$, $P_2 = 7$ in **33 lines, that is 23.4 percent**. The reading adds a single instance to the 32 the lexicon supplies.

The rate is given as it stands. On this measure the corpus is favorable to the accentual model, and the figure is reported without qualification.

The prose baseline removes the prototype's claim to be metrical. In the control corpus, the eight lines whose lexicon yields the sequence 4·4 are realized with $P_1 = 3$ and $P_2 = 7$ in all eight cases. The configuration arises automatically, in ordinary prose, whenever the lexicon supplies two words of four syllables, and it requires no metrical intention. What Condition 6 tests is therefore not whether the configuration exists — it plainly does, in prose as in verse — but whether the poem produces it more often than the lexicon alone would. Here it does not: the reading adds one instance to thirty-two, which is to say it contributes nothing the lexicon had not already provided.

The decisive figure comes from within the accentual model's own successes. Among the 86 lines that satisfy Axiom 1, the two prominences fall on positions 3 and 7 in 33 lines, on 2 and 7 in 30, on 4 and 7 in 19, and on three further configurations in one or two lines each. The *Grenzschema* leads the sequence 3·5 by three lines, 33 against 30. **Fewer than four lines in ten among those that vindicate the accentual school on the number vindicate it on the position**, and the second most frequent configuration places the first prominence on position 2 and not on position 3. In the prose corpus the dispersion is likewise flat, and there too no first position is favored. A prototype realized in fewer than two cases out of five among the cases most favorable to it, and barely ahead of its nearest rival, is not a cognitive attractor.

5.d. Executing Condition 7: the absence of periodicity

Condition 7 is where the accentual model fails without remainder, and it fails on ground that requires no recording at all.

A metrical foot is defined by recurrence: it is a unit that repeats. A line therefore exhibits an internal periodicity only if its accentual nuclei are of equal volume, that is, if $|A_1| = |A_2| = \ldots = |AG|$. The sequence 4·4 satisfies this; 3·5, 5·3 and 2·3·3 do not, since they juxtapose units of different sizes, which no metrical system would call feet.

In this corpus the nuclei of a line are all of equal volume in **32 lines of 141, that is 22.7 percent** — which is to say in the 4·4 lines and in no others. In the remaining 109 there is no repeating unit whatever, and consequently no foot.

The 79 lines whose lexicon yields exactly two groups distribute as follows, every figure being computed from the printed text and depending on no recording, no reader and no ear: the sequence 4·4 in 32 lines, with the stresses on positions 3 and 7 and an interval of four syllables; 3·5 in 25, on positions 2 and 7 with an interval of five; 5·3 in 18, on positions 4 and 7 with an interval of three; and four lines distributed over 4·3, 5·4 and 2·6.

The first stressed syllable falls on position 1, 2, 3 or 4, and the interval separating the two stresses varies from three to six syllables. There is no recurrent unit and no constant interval. Two beats at a variable distance do not constitute a rhythm; they constitute two beats. This result is independent of everything that could reasonably be disputed in this study: even if Axiom 1 were granted in full, even if every line were conceded to carry exactly two accentual groups, no periodic structure would follow and no trochaic foot could be recovered.

Across lines. The second scale of the condition asks whether the poet sequences these configurations, and its criterion is independence: the observed frequency of identical sequences of nucleus volumes in consecutive lines is compared with the frequency expected if each line's configuration were drawn independently, given the distribution of configurations in the corpus. Across the 140 consecutive pairs of the pool, two successive lines carry the same auditory sequence in **25 pairs, that is 17.9 percent**. Independence predicts 22.2 pairs, that is 15.9 percent. The observed frequency exceeds the expected one by fewer than three pairs in a hundred and forty. Within the stanza, where a compositional effect would be most expected, 19 pairs of 102 are identical, that is 18.6 percent. Of the thirty-five stanzas whose four lines are all retained, **1** carries the same sequence on all four. The succession of accentual profiles from one line to the next is what independent draws from the corpus distribution would produce, which is to say that it is governed by the morphological size of the words the poet selected to fill the syllabic count, and by nothing else.

5.e. Executing Condition 8: the tetrasyllabic foot

Condition 8 tests Minerba's derivation of the Swahili line from an imported tetrasyllabic template against the metrical boundaries the corpus imposes. The *utenzi* cannot falsify it: the line counts eight *mizani*, a multiple of four, so a partition into two tetrasyllabic units is arithmetically available in every regular line. The condition is decisive

only where a segment cannot accommodate the template at all, which is the case of the hyper-fractionated *wimbo* matrices examined later in this chapter. What this case study can establish is whether the template, being available, is actually used.

The figure must be given as it stands. A tetrasyllabic nucleus of any kind occurs in **49 lines of 141, that is 34.8 percent**, and **32 lines** are partitioned as 4·4, that is 22.7 percent, and could be described as two tetrasyllabic feet. The remaining 17 lines containing a tetrasyllabic nucleus carry a single one, flanked by units of two or three syllables, in which the four-syllable group is not a repeating unit but the shape of one long word. In 92 lines, that is 65.2 percent, no tetrasyllabic nucleus occurs at all.

The template is available in every line and taken up in fewer than a quarter of them. A metrical unit that the language permits everywhere and that the composer employs in twenty-three lines in a hundred is not the generative principle of the meter.

5.f. Analytical conclusion: the terminal anchor

One constant survives every measurement in this case study, and it is the one the accentual school has consistently mistaken for evidence. Across the 141 lines a nucleus falls on position 7 in every regular line, and it is never unrealized in the reading: not once, in a corpus where nuclei on positions 1, 2, 4 and 5 are lost twelve times between them.

The invariance is real. It carries no metrical information whatever. It follows mechanically from the penultimate rule applied to the last word of the line. Classical Swahili poetics prohibits enjambment, so the line ends on a word boundary; that word carries its lexical stress on its penultimate syllable; and in an eight-syllable line the penultimate of the final word is position 7. Stated lexically rather than positionally — the terminal stress is the penultimate of the last word — the invariant holds without a single exception in this corpus, including in the *guni*, where it falls on position 6 or 8 according to the length of the line. In the prose control corpus the same invariant holds: the second prominence of every two-group line falls on position 7, in sentences that have no meter at all.

Axiom 2 of section 2.5.1 concedes as much, since Vierke derives $P_2 = 7$ from the penultimate stress of the phonological word terminating the line. The axiom therefore states a fact of Swahili morphology and presents it as a metrical law. To take this by-product of word structure, observe that it recurs at the end of every line, and infer from it a trochaic nucleus toward which the verse is built, is to draw the target

around the shot after it has been fired. What remains of the accentual model, once this single invariant is subtracted, is a count that holds in three lines out of five and a placement that holds in fewer than four of every ten among those.

6. Instrumental Verification of the Hostile Sample

Instrumental analysis is applied to a sample of lines and to no more than that, as section 9.2.3 sets out. The exhaustive verification of this corpus is auditory, and no statistic in this case study rests on the measurements reported below. What the sample establishes is that the categories counted in section 5 have measurable physical correlates, and where they do not, that too is reported.

6.a. Composition of the sample

The composition was fixed before any measurement was made. It follows the requirements stated in section 9.2.3, one line for each category the corpus actually produces, together with two lines drawn at random. Two of the categories used in Case Study II have no equivalent here, since no line of this corpus yields one accentual group or four; they are replaced by two categories proper to this poem.

A line conforming exactly to the *Grenzschema*. Stanza 31, line 1, *nachelea ni mjinga*: lexical sequence 4·4, stresses on positions 3 and 7, realized exactly. This is the accentual model's ideal matrix, and it is measured first so that the model is tested where it is strongest. The *Rasi 'lGhuli* produces this configuration in 22.7 percent of its lines, twice the rate of the *Utendi wa Haudaji*, which makes the test more demanding here than there.

An asymmetrical line yielding G-aud = 2. Stanza 28, line 4, *kutunga kwa ushairi:* lexical sequence 3·5, stresses on positions 2 and 7, realized exactly. The count agrees with the bipartite axiom and the positions do not agree with the *Grenzschema*.

A line yielding G-aud = 3. Stanza 31, line 4, *yataka mtu mahiri:* stresses on positions 2, 4 and 7, realized exactly, with the middle nucleus heard as the weaker of the three.

A line carrying a prominence on a monosyllable. Stanza 32, line 1, *wala mi si mualimu*: the lexicon yields 2·6, since the monosyllables *mi* and *si* are proclitic and carry no lexical stress, and the reading realizes 3·1·4, placing a prominence on *si*. The penultimate rule cannot produce a nucleus of one syllable. What the reading records is the oratorical accent, and it is measured here because it is

the clearest instance in either *utenzi* corpus of a prominence that the text does not supply.

A line carrying a nucleus judged weak. Stanza 12, line 4, *hiyo siku ya nushuri*: the reading realizes 2·2·4, with the nucleus on position 1 marked as uncertain. The instrument reports on a decision the ear did not settle, without revising it.

A *guni* of the source text. Stanza 29, line 4, *wajuao tafsiri*: seven syllables, the reader retaining the Arabic closed syllable *taf-* where the assimilated Swahili *tafisiri* would give eight. The line is retained in the pool and counts against the syllabic matrix; it is measured so that the terminal anchor can be observed where it falls on position 6 rather than 7.

Two lines drawn at random. Stanza 7, line 4, *wakuu hata saghiri*, and stanza 22, line 3, *usafi hata ujue*. The draw was made once, on the population of 141 retained lines ordered by stanza and then by line, with a seed derived from the ISBN of this book as described in section 9.2.3. Lines already belonging to the reasoned sample were skipped and the next in the generated sequence taken; no such skip occurred. The two lines obtained are analyzed and published whatever their profile, and no second draw was made.

6.b. What is measured and what is reported

For each line of the sample, three parameters are measured syllable by syllable: fundamental frequency in semitones relative to the speaker's median and corrected for the declination internal to the line, syllable duration, and mean intensity. The analysis range is derived from the speaker's own distribution rather than set by hand, and the conditioning applied to the F0 track is stated in the header of every script.

Each figure carries three readings of the same line. The lexical reading is derived from the printed text by the penultimate rule. The auditory reading is the annotation of section 4. The instrumental reading is established without reference to either: a syllable counts as instrumentally prominent when its detrended F0 mean is a local maximum across the syllables of the line and lies above the declination. That criterion is free to confirm a nucleus the lexicon predicts, to leave one unconfirmed, or to designate a syllable that neither the lexicon nor the ear proposes.

Where the three readings diverge, the divergence is stated in the caption, position by position, and the auditory annotation is not revised. Fundamental frequency is one parameter of three and is not

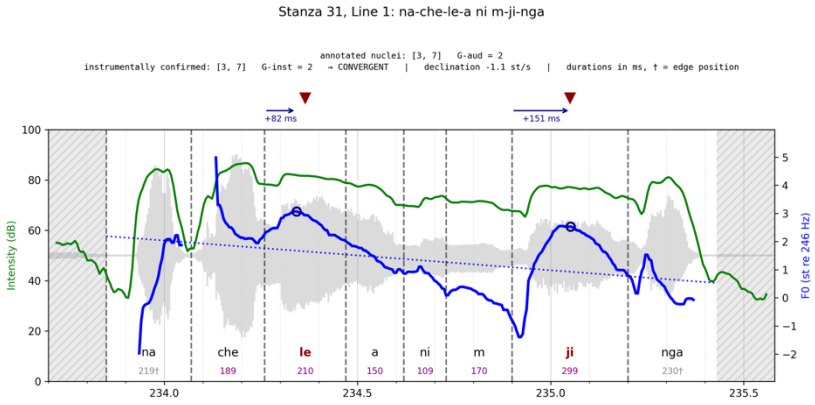
decisive on its own; the durations and intensities are printed beneath every figure whether or not they agree with it.

6.c. Reproduction

One script corresponds to each line of the sample. Each carries in its header the accentual analysis of the line, the boundaries of the line in the recording and the onset of every syllable, and each produces its figure and its caption from the audio alone. The scripts and the recording are deposited as described in section 9.2.7.

6.d. Figures

STANZA 31, LINE 1, NACHELEA NI MJINGA

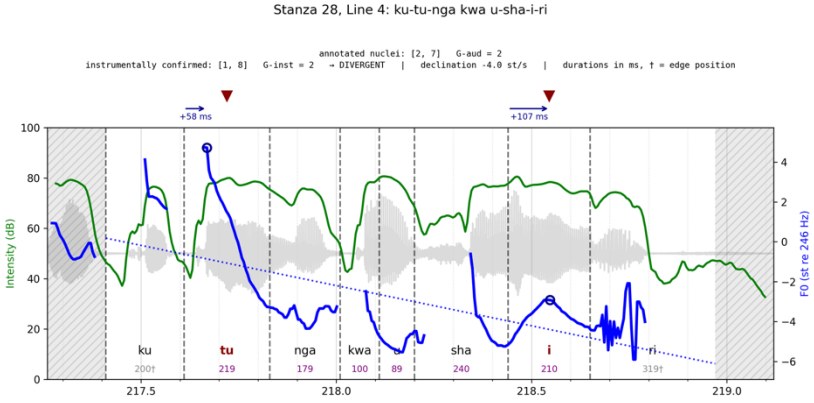



```
Utenzi wa Rasi 'lGhuli, Stanza 31, Line 1: "na-che-le-a ni m-ji-nga"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4·4, P₁=3, P₂=7, G-aud = 2, delta —.
Instrumental reading: 4·4, P₁=3, P₂=7, G-inst = 2, delta — (against
the lexical reading).
Speaker median 246 Hz; declination −1.06 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 na      219 ms†  F0 -1.15 st  peak +184 ms  intensity  62.2 dB
  2 che     189 ms   F0 +0.63 st  peak + 63 ms  intensity  77.6 dB
  3 le      210 ms   F0 +0.97 st  peak + 82 ms  intensity  80.4 dB
  4 a       150 ms   F0 -0.04 st  peak +  1 ms  intensity  75.6 dB
  5 ni      109 ms   F0 -0.56 st  peak + 44 ms  intensity  71.5 dB
  6 m       170 ms   F0 -1.14 st  peak + 43 ms  intensity  70.5 dB
  7 ji      299 ms   F0 +0.52 st  peak +151 ms  intensity  74.6 dB
  8 nga     230 ms†  F0 -0.26 st  peak + 50 ms  intensity  67.4 dB
```

STANZA 28, LINE 4, KUTUNGA KWA USHAIRI

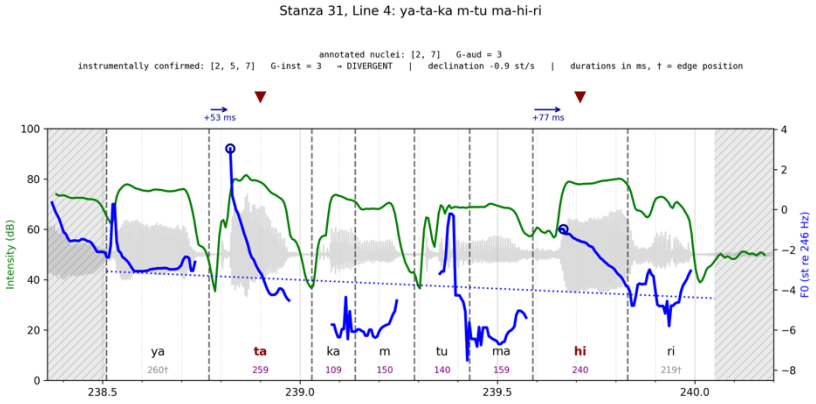



```
Utenzi wa Rasi 'lGhuli, Stanza 28, Line 4: "ku-tu-nga kwa u-sha-i-ri"
Lexical reading: 3·5, P₁=2, P₂=7, G-lex = 2.
Auditory reading: 3·5, P₁=2, P₂=7, G-aud = 2, delta —.
Instrumental reading: 2·6, P₁=1, P₂=8, G-inst = 2, delta −P₁=2, −P₂=7,
+P₁=1, +P₂=8 (against the lexical reading).
Speaker median 246 Hz; declination −4.04 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [2, 7] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [1, 8] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 ku      200 ms†  F0 +2.63 st  peak +100 ms  intensity  57.8 dB
  2 tu      219 ms   F0 +1.28 st  peak + 58 ms  intensity  71.7 dB
  3 nga     179 ms   F0 -1.77 st  peak +171 ms  intensity  69.9 dB
  4 kwa     100 ms   F0 -1.21 st  peak + 65 ms  intensity  62.7 dB
  5 u        89 ms   F0 -2.27 st  peak + 89 ms  intensity  77.2 dB
  6 sha     240 ms   F0 -0.72 st  peak +143 ms  intensity  68.6 dB
  7 i       210 ms   F0 +0.63 st  peak +107 ms  intensity  76.8 dB
  8 ri      319 ms†  F0 +1.03 st  peak +100 ms  intensity  55.8 dB
```

STANZA 31, LINE 4, YATAKA MTU MAHIRI

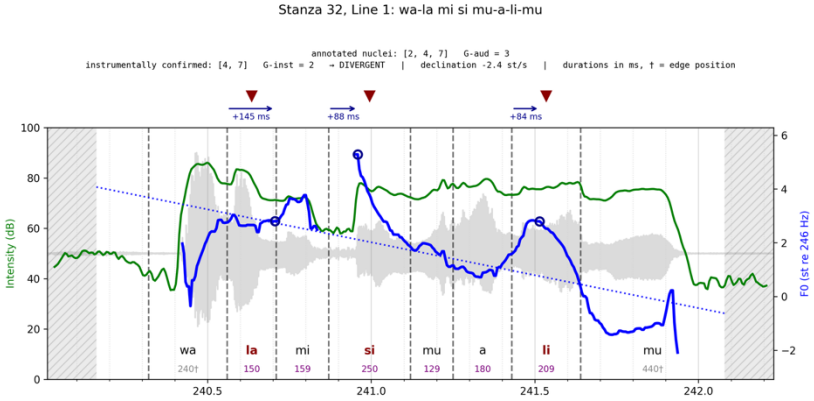



```
Utenzi wa Rasi 'lGhuli, Stanza 31, Line 4: "ya-ta-ka m-tu ma-hi-ri"
Lexical reading: 3·2·3, P₁=2, P₂=4, P₃=7, G-lex = 3.
Auditory reading: 3·2·3, P₁=2, P₂=4?, P₃=7, G-aud = 3, delta –.
Instrumental reading: 3·3·2, P₁=2, P₂=5, P₃=7, G-inst = 3, delta –P₂=4,
+P₂=5 (against the lexical reading).
Speaker median 246 Hz; declination –0.87 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [5] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 ya      260 ms†  F0 +0.56 st  peak + 20 ms  intensity  71.0 dB
  2 ta      259 ms   F0 +0.89 st  peak + 53 ms  intensity  65.6 dB
  3 ka      109 ms   F0 -2.25 st  peak + 86 ms  intensity  65.0 dB
  4 m       150 ms   F0 -2.20 st  peak +105 ms  intensity  62.4 dB
  5 tu      140 ms   F0 +0.51 st  peak + 94 ms  intensity  61.8 dB
  6 ma      159 ms   F0 -2.14 st  peak +127 ms  intensity  66.4 dB
  7 hi      240 ms   F0 +2.02 st  peak + 77 ms  intensity  73.2 dB
  8 ri      219 ms†  F0 +0.15 st  peak +159 ms  intensity  62.8 dB
```

STANZA 32, LINE 1, WALA MI SI MUALIMU

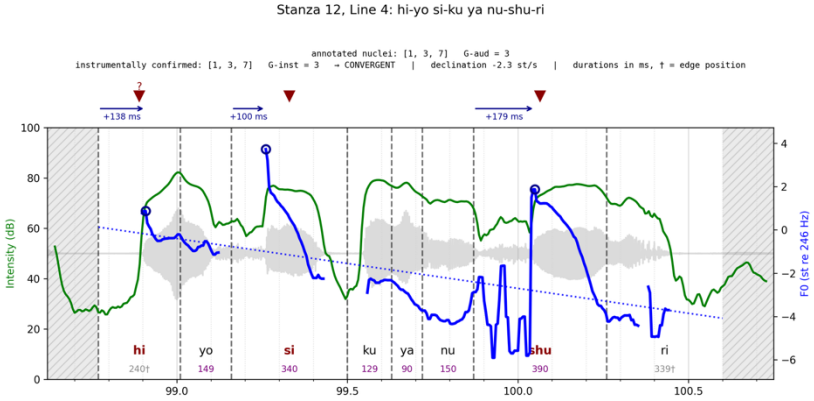



```
Utenzi wa Rasi 'lGhuli, Stanza 32, Line 1: "wa-la mi si mu-a-li-mu"
Lexical reading: 2·6, P₁=1, P₂=7, G-lex = 2.
Auditory reading: 3·1·4, P₁=2, P₂=4, P₃=7, G-aud = 3, delta −P₁=1,
+P₁=2, +P₂=4.
Instrumental reading: 5·3, P₁=4, P₂=7, G-inst = 2, delta −P₁=1, +P₁=4
(against the lexical reading).
Speaker median 246 Hz; declination −2.45 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [2] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 wa      240 ms†  F0 -1.52 st  peak +221 ms  intensity  65.0 dB
  2 la      150 ms   F0 -0.17 st  peak +145 ms  intensity  77.1 dB
  3 mi      159 ms   F0 +0.66 st  peak + 89 ms  intensity  68.0 dB
  4 si      250 ms   F0 +1.18 st  peak + 88 ms  intensity  69.8 dB
  5 mu      129 ms   F0 +0.11 st  peak + 42 ms  intensity  74.6 dB
  6 a       180 ms   F0 -0.18 st  peak +175 ms  intensity  76.4 dB
  7 li      209 ms   F0 +1.35 st  peak + 84 ms  intensity  76.3 dB
  8 mu      440 ms†  F0 -1.07 st  peak +281 ms  intensity  63.1 dB
```

STANZA 12, LINE 4, HIYO SIKU YA NUSHURI

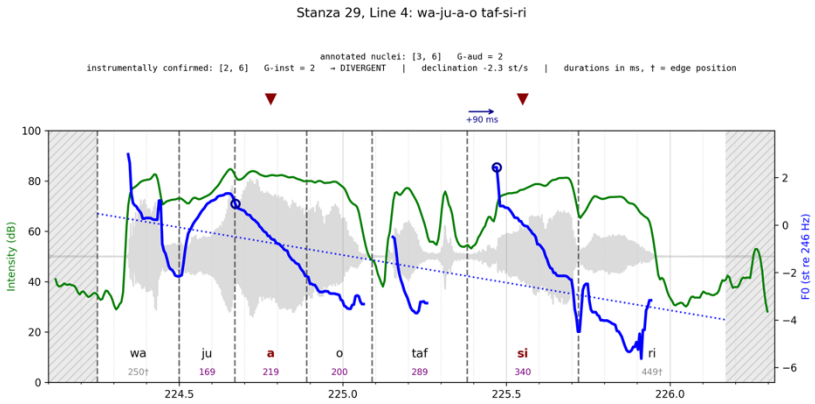



```
Utenzi wa Rasi 'lGhuli, Stanza 12, Line 4: "hi-yo si-ku ya nu-shu-ri"
Lexical reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-lex = 3.
Auditory reading: 2·2·4, P₁=1?, P₂=3, P₃=7, G-aud = 3, delta —.
Instrumental reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-inst = 3, delta —
(against the lexical reading).
Speaker median 246 Hz; declination −2.30 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 hi      240 ms†  F0 +0.15 st  peak +138 ms  intensity  52.0 dB
  2 yo      149 ms   F0 -0.24 st  peak +  2 ms  intensity  70.6 dB
  3 si      340 ms   F0 +1.16 st  peak +100 ms  intensity  62.9 dB
  4 ku      129 ms   F0 -0.62 st  peak +127 ms  intensity  65.1 dB
  5 ya       90 ms   F0 -1.06 st  peak +  2 ms  intensity  76.1 dB
  6 nu      150 ms   F0 -1.78 st  peak +146 ms  intensity  70.7 dB
  7 shu     390 ms   F0 +0.80 st  peak +179 ms  intensity  69.3 dB
  8 ri      339 ms†  F0 -0.63 st  peak +122 ms  intensity  52.7 dB
```

STANZA 29, LINE 4, WAJUAO TAFSIRI



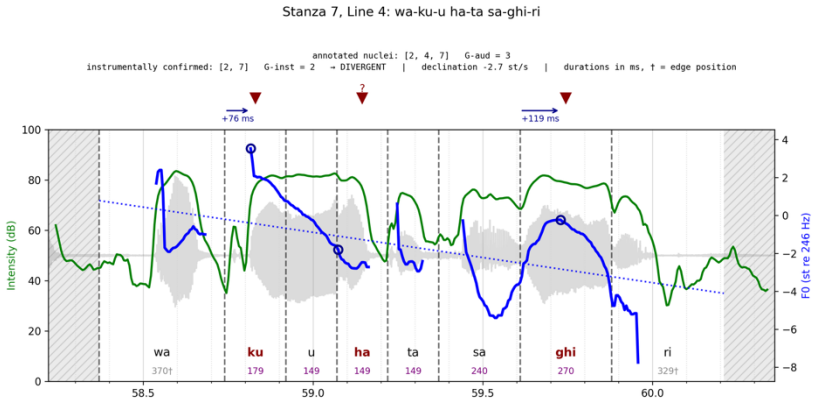
Utenzi wa Rasi 'lGhuli, Stanza 29, Line 4: "wa-ju-a-o taf-si-ri"
Lexical reading: 4·3, P_1=3, P_2=6, G-lex = 2.
Auditory reading: 4·3, P_1=3, P_2=6, G-aud = 2, delta —.
Instrumental reading: 3·4, P_1=2, P_2=6, G-inst = 2, delta $-P_1$=3, $+P_1$=2 (against the lexical reading).
Speaker median 246 Hz; declination –2.33 st/s.

The instrumental reading diverges from the auditory reading: position(s) [3] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	wa	250 ms†	F0 -0.16 st	peak + 94 ms	intensity	61.7 dB
2	ju	169 ms	F0 +0.75 st	peak +151 ms	intensity	76.2 dB
3	a	219 ms	F0 +0.21 st	peak + 1 ms	intensity	81.8 dB
4	o	200 ms	F0 -1.76 st	peak + 4 ms	intensity	73.1 dB
5	taf	289 ms	F0 -1.02 st	peak + 62 ms	intensity	62.3 dB
6	si	340 ms	F0 +1.86 st	peak + 90 ms	intensity	72.7 dB
7	ri	449 ms†	F0 -0.95 st	peak + 28 ms	intensity	53.5 dB

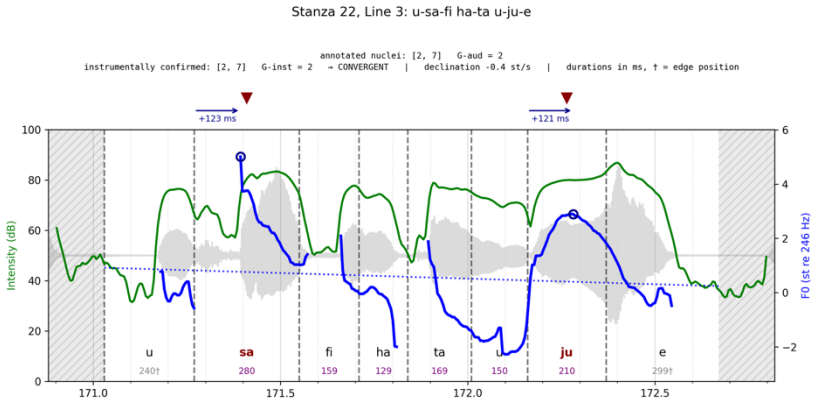
STANZA 7, LINE 4, WAKUU HATA SAGHIRI




```
Utenzi wa Rasi 'lGhuli, Stanza 7, Line 4: "wa-ku-u ha-ta sa-ghi-ri"
Lexical reading: 3·2·3, P₁=2, P₂=4, P₃=7, G-lex = 3.
Auditory reading: 3·2·3, P₁=2, P₂=4?, P₃=7, G-aud = 3, delta —.
Instrumental reading: 3·5, P₁=2, P₂=7, G-inst = 2, delta −P₂=4 (against
the lexical reading).
Speaker median 246 Hz; declination −2.66 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [4] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 wa     370 ms†  F0 -0.85 st  peak +183 ms  intensity  57.8 dB
  2 ku     179 ms   F0 +2.31 st  peak + 76 ms  intensity  68.7 dB
  3 u      149 ms   F0 +0.53 st  peak +  5 ms  intensity  81.7 dB
  4 ha     149 ms   F0 -1.33 st  peak +  4 ms  intensity  68.8 dB
  5 ta     149 ms   F0 -0.62 st  peak + 27 ms  intensity  64.6 dB
  6 sa     240 ms   F0 -1.76 st  peak + 71 ms  intensity  68.7 dB
  7 ghi    270 ms   F0 +1.49 st  peak +119 ms  intensity  78.0 dB
  8 ri     329 ms†  F0 -1.65 st  peak + 17 ms  intensity  51.4 dB
```




```
Utenzi wa Rasi 'lGhuli, Stanza 22, Line 3: "u-sa-fi ha-ta u-ju-e"
Lexical reading: 3·2·3, P₁=2, P₂=4, P₃=7, G-lex = 3.
Auditory reading: 3·5, P₁=2, P₂=7, G-aud = 2, delta −P₂=4.
Instrumental reading: 3·5, P₁=2, P₂=7, G-inst = 2, delta −P₂=4 (against
the lexical reading).
Speaker median 246 Hz; declination −0.41 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 u      240 ms†  F0 -0.80 st  peak +155 ms  intensity  53.8 dB
  2 sa     280 ms   F0 +1.65 st  peak +123 ms  intensity  73.4 dB
  3 fi     159 ms   F0 +0.05 st  peak +111 ms  intensity  61.9 dB
  4 ha     129 ms   F0 -0.90 st  peak + 35 ms  intensity  68.2 dB
  5 ta     169 ms   F0 -1.00 st  peak + 54 ms  intensity  69.7 dB
  6 u      150 ms   F0 -2.30 st  peak +147 ms  intensity  71.3 dB
  7 ju     210 ms   F0 +1.77 st  peak +121 ms  intensity  77.9 dB
  8 e      299 ms†  F0 -0.02 st  peak +  1 ms  intensity  63.4 dB
```

6.e. The concordance of the instrumental reading

The protocol of section 9.2.3 requires that this step report a concordance rate together with its failures. Of the eight lines measured, the instrumental reading coincides exactly with the auditory annotation in **three** — 31.1, 12.4 and 22.3 — and departs from it in **five**. The rate is therefore three of eight, and the departures are set out below rather than summarized away.

The second of these, 12.4, *hiyo siku ya nushuri*, deserves a sentence of its own: lexicon, ear and instrument converge on three accentual

groups, 2·2·4, at positions 1, 3 and 7, an eight-syllable line carrying the configuration the bipartite axiom declares impossible, confirmed by every procedure this study employs.

The confidence mark is vindicated. Three lines of the sample carry a nucleus that the ear heard but judged weak: 12.4 on position 1, 7.4 and 31.4 on position 4. The first is confirmed; in the two others, the nucleus is not. In neither is that nucleus confirmed. At 7.4 the instrument drops it altogether, the syllable *ha* falling to −1.33 semitones below the declination; at 31.4 it moves it to position 5, *m* measuring −2.20 and *tu* +0.51. Where the listener hesitated, the instrument found no clean prominence in two cases of three. The mark introduced in section 9.2.3 is therefore not a rhetorical precaution: it identifies in advance the positions on which two independent procedures fail to agree.

The terminal anchor holds under measurement. The instrument confirms a prominence on position 7 in six of the eight lines, and on position 6 in the seven-syllable *guni* at 29.4, where the penultimate of the final word falls there. The single exception is 28.4, discussed below. The invariant established on the printed text in section 5.f survives the transfer to the signal.

Two departures are displacements of a single syllable, and two are omissions. At 31.4 the instrument places the middle prominence on position 5 where the ear places it on 4; at 29.4 it places the first on position 2 where the ear places it on 3; at 32.1 and 7.4 it declines to confirm a position the ear proposed without proposing another in its place. The two displacements do not run the same way, and neither follows from a property of the meter. At 31.4 the ear's syllable is the syllabic nasal *m*, which carries no vowel, and the pitch movement is realized on the vowel that follows. At 29.4 the instrument moves one syllable earlier, to *ju*, although the syllable the ear chose, *a*, is both the longer and the louder of the two (219 milliseconds against 169, 81.8 dB against 76.2). The contour, read syllable by syllable, follows pitch alone; the ear weighs pitch, length and loudness together.

One departure reverses the reading entirely, and it is the most instructive line of the sample. At 28.4, *kutunga kwa ushairi*, the ear places the two prominences on positions 2 and 7. The instrument places them on positions **1 and 8** — the two edge syllables. The line carries the steepest declination of the sample, −4.04 semitones per second, and once that trend is removed the highest detrended values fall on the line-initial reset, *ku* at +2.63, and on the

final boundary rise, *ri* at +1.03, while the syllable the ear hears as prominent, *tu*, measures +1.28 and the terminal nucleus *i* only +0.63. What the instrument records here is not the accentual structure of the line at all: it is the pitch reset that opens a line of verse and the boundary tone that closes it.

What this establishes about the two models.

For the syllabic model, nothing. The syllable count and the rhyme boundaries are properties of the printed text, and no measurement of prominence can confirm or threaten them. That the instrumental reading varies from line to line leaves Conditions 1 and 2 exactly where section 5.a left them: N = 8 verified in 97.9 percent of the pool, the global rhyme X^s holding in forty stanzas of forty, the local rhyme a^s in thirty-eight. The syllabic matrix is indifferent to this variation because it never asks where the prominence falls. It is worth noting, as an empirical corollary, that this syllabic segmentation is visually corroborated across the sample figures by the acoustic intensity track (plotted in green). The successive energetic peaks consistently align with the syllabic nuclei, demonstrating the physical reality of the absolute syllable count independently of the fluctuating Fo contour.

For the accentual model, a great deal, and the point is not that the instrument contradicts the ear. It is that both readings are readings of the same performance, obtained by procedures each of which the accentual tradition would accept, and they disagree in five lines of eight. Locate the beats by ear and they fall on positions 2 and 7 at 28.4; locate them by pitch alone, as an instrument does, and they fall on positions 1 and 8. A meter that placed a beat on the final syllable of every line would be unrecognizable, and no accentual theory of Swahili has ever proposed one.

There is therefore no single fact of the matter about where the beats of a line lie, even within one reading, by one speaker, measured one way. The syllable count admits no such indeterminacy: it is eight, the signal shows eight nuclei, and eight is what any reader counts on the page. A theory that takes prominence for the generative principle of the meter builds on a quantity that has no stable value; a theory that takes the syllable builds on one that has nothing else.

From the perspective of Occam's razor (*lex parsimoniae*), the disparity forces a final question. On one side stand endogenous syllabic principles, verified here at 97.9 percent for the syllable count and without exception for the global rhyme, and documented in endogenous terminology by the poets themselves as the foundation of their art, as Chapter 4 establishes. On the other stands an exogenous

accentual matrix resting on a *Grenzschema* that the lexicon supplies rather than the meter, and that survives only through a prescriptive reduction ignoring alternative configurations of nearly equal frequency — the sequence 3·5 trails it by three lines in this corpus. Given this disparity, must we not simply apply Occam's razor and excise the accentual school's theoretical superstructure entirely from Swahili metrical analysis?

9.4.2. Case Study II: *Utendi wa Haudaji* — Unknown Author (19th Century)

To rigorously test the explanatory power of the endogenous syllabic metrical grid against the exogenous hypotheses of the accentual school, both formal models—the formalized syllabic matrix alongside its corresponding algebraic equation (see QRG 2.), and Vierke's formalized accentual model (see Section 2.5.1.)—are tested directly against the lexical and auditory readings of the corpus, strictly in accordance with the eight conditions of falsification defined in Section 9.2.4. The text subjected to this experimental verification is the *Utendi wa Haudaji* (The Poem of the Palanquin) (Vierke, 2011). While the precise date of its composition and the exact identity of its author remain lost to history, the poem was composed on the island of Lamu during the 19th century.

It is essential to first acknowledge and salute the monumental philological achievement of Clarissa Vierke in editing and establishing this pristine text from disparate and complex manuscript sources, even as we maintain the historiographical reservations expressed in Chapter 1 regarding the neutral, uncritical contextualization of the Ernst Dammann archive upon which part of her edition relies. Her meticulous textual reconstruction provides an invaluable foundation for Swahili studies. However, while the philological establishment of the text is impeccable, the metrical theory she subsequently projects onto it must be subjected to strict scientific tests.

Testing the accentual model against this specific 19th-century text holds paramount epistemological importance: this poem constitutes the exact corpus from which Vierke derived the foundational postulates of the contemporary accentual school, specifically the concept of the *Grenzschema* (see Section 2.3.1). By objectively evaluating this theoretical model against the lexical, auditory, and acoustic reality of its own foundational text, strictly guided by our falsification conditions, we will determine whether the framework

that was later projected onto other classical Swahili meters actually withstands empirical scrutiny.

1. Philological Context and Methodological Protocol

The *Utendi wa Haudaji* relies on a tripartite manuscript tradition, reflecting the fluid and dynamic nature of Swahili textual transmission across generations. The primary sources include:

- Manuscript DA (Dammann Collection): A comprehensive 914-stanza transliterated in 1936/37 by Ernst Dammann and the Swahili scholar Muhamadi Kijuma.
- Manuscript A1 (Allen Collection): A 34-page manuscript in Arabic script, historically containing omissions that reduce its actual length to 661 stanzas.
- Manuscript A2 (Allen Collection): Titled *Utendi wa Hijira*, an 826-stanza manuscript in Arabic script.

The systematic cross-referencing of Dammann's transliteration (DA) with the independent Arabic script manuscripts from the Allen collection (A1, A2) allows for a robust, triangulated textual foundation.

Our transcription of the text, which served as the reading sheet for the experiment, reflects the phonological specificities of the Kiamu dialect, such as the aspiration of voiceless consonants and the characteristic rhotacism of nasal compounds like /ndr/. The diacritical markings originally distinguishing dental plosives (t̠, d̠, nd̠) in earlier philological transcriptions were systematically omitted in the reading sheet to avoid visual clutter and prevent reading artifacts. Apart from this editorial simplification, the dialectal forms were left intact in the text supplied to the reader, which was handed over without further graphic normalization. What the recording documents is therefore her own adaptation of them: she reads *ndiyo* for *ndriyo*, *mtu* for *mthu*, *zote* for *zothe*, applying the macro-Swahili phonological competence she commands to a nineteenth-century Kiamu text she does not practice. Every such divergence between the supplied text and the execution is listed in the divergence table below, together with its effect on the syllable count. None of these substitutions alters that count, and none displaces a stress: *ndri-o* and *ndi-o* are alike disyllabic, as are *m-thu* and *m-tu*. The dialectal modernization is metrically inert, and the mapping is unaffected by it. The text supplied to the reader is deposited unchanged in the Zenodo

repository alongside the recordings, so that every divergence recorded below can be verified against the stimulus she actually received.

This analysis is based on a formal reading by our L1 Swahili speaker, which sets aside the melodic execution of Layer 3 and records how much of the lexical accentual structure a reader actually realizes. This anchors our analysis in the text as composed (see Section 3.3.). Layer 1 and Layer 2 constitute the fundamental syllabic law and the architecture of the text, and they are present in this study as a material object: the page handed to the reader, which carries the syllabic count, the rhyme boundaries, the lines, the stanzas, and nothing else. Upon it the language imposes its own accentual structure, deterministically, through the fixed penultimate stress of the phonological word. It is imperative to oppose both of these deterministic frameworks—the formal meter based on syllabic principles and the linguistic stress—to Layer 3, which encompasses the musical execution, the melody (*mahadhi*), and the performative chanting of the poem.

To guarantee reproducibility and to avoid any Procrustean reductionism, the syllabification and the accentual mapping of this text follow the three-step protocol set out in Section 9.2.3. The accentual structure of each line is first derived from the modernized phonetic sequence—systematically accounting for the reader's dialectal adaptations—by applying the penultimate rule, an operation that requires no recording and that any reader may repeat. The formal reading is then screened line by line by ear, and the two results are reported side by side for all 160 lines. Acoustic extraction is applied to a sample of lines only, and no statistic in this case study rests upon it.

For the lines of that sample, the instrumental reading algorithmically designates prominence based strictly on the fundamental frequency (F0) contour, as F0 is the standard acoustic parameter traditionally used in Swahili linguistics to measure accentual peaks. Syllabic duration and acoustic intensity are simultaneously extracted to expose the full empirical footprint of the recitation. The temporal boundaries of the syllable nuclei are established in Praat, and the acoustic data is processed and plotted through a custom Python script using the Parselmouth library. The specific calculation parameters—such as the speaker median reference and the subtraction of internal declination—are transparently detailed in the caption accompanying each figure. Furthermore, the complete script used to evaluate these accentual positions is openly accessible in the project's GitHub repository. This double transparency allows any researcher to read the exact

methodology alongside the visual data, perfectly reproduce the results, or modify the code to test alternative hypotheses.

For this empirical verification, the exhaustive phonetic mapping of the first 40 stanzas will be systematically analyzed to explicitly illustrate the structural dynamics of the verse and draw our statistical conclusions regarding the positional accentual matrix (such as P_1 and P_2) across the corpus.

2. Identification of the Formal Syllabic Matrix

Before proceeding with the empirical verification, we must formally define the specific endogenous model that will be subjected to testing. The structural architecture of the *utenzi* meter is formally defined by a strict syllabic matrix. The stanza (*ubeti*) is a highly stable architectural block composed of four lines. This form dictates that each line contains exactly 8 syllables (*mizani*). The first three lines share a rhyme (*a*) whose scope is strictly limited to the individual stanza, changing from one stanza to the next, while the fourth line carries the global rhyme (*bahari*, denoted X) that spans the entirety of the poem, locking it into a cohesive structural unity. Both variables are written unresolved in the matrix, and each is governed by the hierarchy defined in the symbolic alphabet, $a^s > a^v$ and $X^s > X^v$: the syllabic form is the rule and is tested first, the vocalic form a license examined only where the syllabic one has been contradicted. Drawing upon the theoretical formalization presented in Chapter 8 (specifically Section 8.3 on the *utenzi* meter) and summarized in the Quick Reference Guide (QRG.2), the formal generative matrix selected for this experimental evaluation operates as follows:

Formal Parameters:
$n_{max} = 4, S = 8$

The Formal Matrix:
L1: [8, a] //
L2: [8, a] //
L3: [8, a] //
L4: [8, X] //

The Algebraic Equation:
$\boldsymbol{Utenzi} = \mathbf{3}(\boldsymbol{N_a}//) \cdot \boldsymbol{N_X}//$ (where N=8)

3. The Competing Generative Model: Vierke's Grenzschema

The accentual school superimposes a distinct, exogenous accentual model upon this pristine syllabic grid. Clarissa Vierke postulates that the generative engine of the *utenzi* is not the absolute syllable count, but an underlying rhythmic beat. Her model asserts that the 8-syllable line is inherently structured by two trochaic feet ($G = 2$). Specifically, she defines the *Grenzschema* as a cognitive metrical template that rigidly anchors the primary stresses on the exact third and seventh syllables of the line ($P_1 = 3$ and $P_2 = 7$).

In this theoretical paradigm, the morphological complexity of the Swahili verse is reduced to a simplistic, binary rhythmic predictability centered on tetrasyllabic (4-4) symmetry. If this hypothesis holds as a generative law, then the accentual structure of the first 40 stanzas must confirm the predominance of the bipartite configuration with nuclei on P_1=3 and P_2=7. The test is applied twice, as Condition 5 requires: to the lexicon of each line, which no recording is needed to establish, and to what the formal reading realizes of it.

4. Utendi wa Haudaji: *Exhaustive Syllabic and Accentual Mapping, Stanzas 1–40*

4.a. Reading the entries

Each entry corresponds to a poetic line. It occupies a block of three short lines, in the notation of sections QRG.1 and 2.5. The first line gives the reference, the syllabified text, and its syllable count; *guni* marks a poetic line whose count departs from eight in the source text itself.

The second line is the lexical reading, derived from the printed text by the penultimate rule and requiring no recording: the sequence of nucleus volumes |A1|·|A2|·…, the positional index P*i* of the stressed syllable of each nucleus A*i*, and G-lex. Since the stressed syllable of a nucleus of |A*i*| syllables is its syllable |A*i*| − 1, the sequence of volumes determines the positions.

The third line is the auditory reading, established by ear on the formal reading, followed by Δ, which records the transition from one to the other. The notation −P*i*=n marks the lexical nucleus A*i*, whose stress falls on syllable n, as unrealized; +P*i*=n marks a realized nucleus A*i*, with its stress on syllable n, that the lexicon does not predict; a dash marks a realization that follows the lexical structure exactly.

Where the poetic line is excluded from the statistical pool, this short line gives the ground of exclusion instead.

A question mark on a positional index marks an auditory decision made without full confidence: P2=5? is a nucleus heard but judged weak and counted nonetheless. The position is counted as written, and the mark registers the doubt without resolving it. No such mark can appear in the lexical reading, which is derived by rule.

4.b. Mapping

Stanza 1
1.1 bis-mi-lla-hi a-wa-li (7)
lexical 4·3 — P1=3, P2=6 — G-lex 2
excluded — Arabized execution
1.2 i-si-mu ya-ke ja-la-li (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
1.3 ndi-o mwa-ndo wa ka-u-li (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ –P1=1
1.4 wa m-tu ku-ta-m-ki-ya (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 2
2.1 a-rah-ma-ni a-zi-zi (7)
lexical 4·3 — P1=3, P2=6 — G-lex 2
excluded — Arabized execution
2.2 na-bi ndi-ye mu-a-wa-zi (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ –P1=1
2.3 we-ma wa-wi hu-wa-ja-zi (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ –P3=5
2.4 u-li-mwe-ngu-ni ja-li-a (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
Stanza 3
3.1 na a-ra-hi-mi ka-ri-ma (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —

3.2 ni m-te-u-zi wa we-ma (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
3.3 ke-sho ja-na-ti na-ki-ma (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
3.4 mo-la ta-i-te-u-li-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 4
4.1 al-ham-du li-lla-hi (6)
lexical 3·3 — P1=2, P2=5 — G-lex 2
excluded — Arabized execution
4.2 ku-la mwe-nye ku-ma-di-hi (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ −P1=1
4.3 nda-kwe a-si-yo-sha-bi-hi (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 8 — P1=7 — G-aud 1 — Δ −P1=1
4.4 mo-la wa zi-u-mbe pi-a (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·4·2 — P1=1, P2=5, P3=7 — G-aud 3 — Δ —
Stanza 5
5.1 si-fa zo-te nje-ma hi-zi (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5?, P4=7 — G-aud 4 — Δ —
5.2 ze-ne-o ka-ma mwe-zi (7) — *guni*
lexical 3·2·2 — P1=2, P2=4, P3=6 — G-lex 3
auditory 3·2·2 — P1=2, P2=4, P3=6 — G-aud 3 — Δ —
5.3 ni za-ke ye-ye a-zi-zi (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ −P2=4
5.4 ha-zi-ka-di-ri m-mo-ya (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ +P1=1
Stanza 6
6.1 na ku-to-sha zi-a-da-ye (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

6.2 ni re-he-ma ya-ke ye-ye (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ −P2=5
6.3 mwe-nye e-zi mu-ngu ndi-ye (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ −P1=1, −P3=5
6.4 a-mba-o ka-o-ne-ke-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 7
7.1 na wo-ko-fu-we ka-mi-li (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
7.2 hu-mwe-nde-le-a mu-ri-sa-li (9)
lexical 5·4 — P1=4, P2=8 — G-lex 2
excluded — *guni* of reading
7.3 na sa-ha-ba-ze na a-li (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 4·1·3 — P1=3, P2=4, P3=7 — G-aud 3 — Δ +P1=3
7.4 mi-le-le hu-wa-shu-ki-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 8
8.1 u-li-ne-ne-le ha-shi-ma (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
8.2 m-tu-me wa-kwe ka-ri-ma (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ −P2=4
8.3 wo-ko-fu-we na re-he-ma (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
8.4 wa mu-ngu hu-tu-i-li-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 9
9.1 tu-mwa we-tu m-fa-dha-li (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —

9.2 ndi-ye mu-u-mbwa a-wa-li (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ −P1=1
9.3 ka-po-wa na ma-na-zi-li (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
9.4 ya-si-yo-m-ti-ndi-ki-a (8)
lexical 8 — P1=7 — G-lex 1
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ +P1=3
Stanza 10
10.1 hi-fa-dhi tu-zi-e kwe-nu (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
10.2 ni me-ngi ni nyi-ngi m-no (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ −P2=5
10.3 ndi-po kha-ko-sa ma-ne-no (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 8 — P1=7 — G-aud 1 — Δ −P1=1, −P2=4
10.4 si-zi-we-zi ku-zi-ta-ya (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 11
11.1 di-ba-ji ni-me-khi-ti-mu (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
11.2 kwa che-o na ta-ka-dhi-mu (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
11.3 sa-sa ta-ka ku-nu-dhu-mu (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —
11.4 ni-me-zo ya-ku-su-di-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 12
12.1 ha-di-thi ni-me-i-o-na (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —

12.2 ya mu-ha-ma-di a-mi-na (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ −P1=4, +P1=3
12.3 ya-li-po-in-gi-a fi-ti-na (9) — *guni*
lexical 6·3 — P1=5, P2=8 — G-lex 2
auditory 6·3 — P1=5, P2=8 — G-aud 2 — Δ —
12.4 ya-kwe na u-ju-ha-li-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 13
13.1 na fi-ti-na zi-ki-kwa-ka (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
13.2 zi-ka-tu-pa na mi-pa-ka (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
13.3 al-ha-ji-ri ma-ka (6)
lexical 4·2 — P1=3, P2=5 — G-lex 2
excluded — Arabized execution
13.4 m-tu-me we-tu na-bi-a (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
Stanza 14
14.1 ma-ka a-li-po-ha-ji-ri (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
14.2 kwe-nda ma-di-na ba-shi-ri (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1?, P2=4, P3=7 — G-aud 3 — Δ —
14.3 we-nye-we wa-li-sha-ki-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
14.4 kwa tu-mwa ku-wa-ngi-li-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 15
15.1 a-li-to-ka ye m-pwe-ke (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

15.2 ha-ku-twa-a ta-a mba-ke (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
15.3 fa-tu-ma na wa-ke wa-ke (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5?, P3=7 — G-aud 3 — Δ —
15.4 ma-ka wa-ka-i-ka-li-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 16
16.1 wa-ke wo-te na zi-ja-na (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
16.2 ha-kwe-nda na-o a-mi-na (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
16.3 na u-te-ndi la-kwe i-na (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
16.4 ya ha-u-da-ji si-ki-a (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
Stanza 17
17.1 u-te-ndi mbwa ha-u-da-ji (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
17.2 ndi-lo i-na-le wa-mbu-ji (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
17.3 za-li-zo-ngi-a si-ra-ji (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
17.4 mu-ha-mma-di ku-mwe-nde-a (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
Stanza 18
18.1 tu-mwa ku-le ma-di-na-ti (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —

18.2 a-li-po-kwi-sa ku-ke-ti (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
18.3 a-li-a-nga-ta sa-u-ti (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
18.4 kamwamkuwa alia
lexical 5·3 — P1=4, P2=7 — G-lex 2
excluded — inaudible execution
Stanza 19
19.1 ka-mba bu-nu a-mu ya-ngu (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
19.2 ku-ru-bi-a ha-pa pa-ngu (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
19.3 ni-ne-ne ma-ne-no ya-ngu (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
19.4 u-pa-te ku-ya-si-ki-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 20
20.1 a-ka-ji-bu ha-i-da-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
20.2 u-pe-nda-lo ni-khu-bi-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
20.3 am-ba-lo wa-li-khi-ta-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
20.4 si-che ku-ni-ka-ri-fi-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 21
21.1 a-ka-ta-m-ka a-mi-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —

21.2 ka-mba a-ba al-ha-sa-ni (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
excluded — Arabized execution
21.3 m-no ni-me-wa-ta-ma-ni (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
21.4 wa-ke wa-ngu na dhu-ri-a (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1?, P2=3, P3=7 — G-aud 3 — Δ —
Stanza 22
22.1 u-si-ku si-la-li te-na (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5?, P3=7 — G-aud 3 — Δ —
22.2 si-o-ni ra-ha m-ta-na (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
22.3 mi na-o ku-to-o-na-na (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
22.4 m-no ya-me-ni-ku-li-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 23
23.1 ni-lo na-lo kwa-mbi-e (7)
lexical 2·2·3 — P1=1, P2=3, P3=6 — G-lex 3
excluded — *guni* of reading
23.2 e-ne-nda ka-ni-twa-li-e (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
23.3 mo-yo wa-ngu u-pu-mba-e (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
23.4 ma-to ki-wa-a-nga-li-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 24
24.1 a-ka-ji-bu ha-i-da-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

24.2 ka-mba e-we mu-khu-ta-ri (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
24.3 ku-tu-mi-ka ni-me-ki-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
24.4 ta-kwe-nda ku-ku-twa-li-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 25
25.1 sa-sa hi-ya ta-to-ka (7)
lexical 2·2·3 — P1=1, P2=3, P3=6 — G-lex 3
excluded — *guni* of reading
25.2 tu-mwa wa-ngu ne-nde ma-ka (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5?, P4=7 — G-aud 4 — Δ —
25.3 ha-o u-me-zo-wa-ta-ka (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
25.4 in-sha-llah ta-kwe-te-a (7) — *guni*
lexical 3·4 — P1=2, P2=6 — G-lex 2
auditory 3·4 — P1=2, P2=6 — G-aud 2 — Δ —
Stanza 26
26.1 ha-yo a-li-po-pu-li-ka (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
26.2 a-li-i me-zo-ta-mu-ka (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
26.3 mo-yo wa-li-m-ku-ndu-ka (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
26.4 na ra-ha i-ka-m-ngi-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 27
27.1 ka-fu-ra-hi mu-u-nga-mu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

27.2 ka-mba e-we bu-nu a-mu (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
27.3 ta-ku-pa na ma-re ma-ka-mu (9) — *guni*
lexical 3·3·3 — P1=2, P2=5, P3=8 — G-lex 3
auditory 3·3·3 — P1=2, P2=5, P3=8 — G-aud 3 — Δ —
27.4 wa-ra-ka wa a-fi-a (7) — *guni*
lexical 3·4 — P1=2, P2=6 — G-lex 2
auditory 3·4 — P1=2, P2=6 — G-aud 2 — Δ —
Stanza 28
28.1 wa-ra-ka mwe-nye zi-a-si (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
28.2 we-nde na-o m-ma-ri-si (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
28.3 kwa ba-ba we-tu u-pe-si (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
28.4 bu-nu a-bi ta-li-bi-a (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
Stanza 29
29.1 tha-kwa-ndi-ki-a wa-ra-ka (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
29.2 we-nde na-o ha-ta ma-ka (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1?, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
29.3 kwa ba-ba we-tu pu-li-ka (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
29.4 a-ba-si mwe-nye ma-fa-a (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 30
30.1 na wa-ra-ka ho-no nda-ni (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —

30.2 ni-ti-zi-e-zo ya-ki-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
30.3 ni wa-ki-a kha-mu-si-ni (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
30.4 mo-ya ha-i-ku-pu-ngu-wa (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 31
31.1 wa-ki-a hi-zo sa-hi-bu (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
31.2 kham-si-ni kwa he-sa-bu (7)
lexical 3·4 — P1=2, P2=6 — G-lex 2
excluded — Arabized execution
31.3 nda fe-dha-ti na dha-ha-bu (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 1·3·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ +P1=1
31.4 a-mu-yo ku-m-ni-khi-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 32
32.1 ba-si u-ye u-ji-li-si (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
32.2 u-i-shi-ke ka-ra-ta-si (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
32.3 ni-kwe-le-ze kwa zi-a-si (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
32.4 ku-la ya-mbo mo-ya-mo-ya (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
Stanza 33
33.1 wa-ra-ka hu-nu na-ndi-ka (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —

33.2 i-li ku-u-pe-ka ma-ka (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·4·2 — P1=1, P2=5, P3=7 — G-aud 3 — Δ —
33.3 wa-pa-te ku-ya-pu-li-ka (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
33.4 ma-ne-no ta-ko-kwa-mbi-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 34
34.1 bis-mi-lla-hi ra-di-di (7)
lexical 4·3 — P1=3, P2=6 — G-lex 2
excluded — Arabized execution
34.2 i-si-mu ya-ke wa-du-di (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4?, P3=7 — G-aud 3 — Δ —
34.3 ndi-mi tu-mwa mu-ha-mma-di (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
34.4 mwe-nye ku-wa-shu-fa-ki-a (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
Stanza 35
35.1 wa-ra-ka u-to-ka kwa-ngu (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5?, P3=7 — G-aud 3 — Δ —
35.2 mi-mi m-tu-me wa mu-ngu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
35.3 ku-wa-kho-fi-sha wa-zu-ngu (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
35.4 na me-ma ku-wa-ke-dhe-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 36
36.1 u-to-ka kwa-ngu ha-shi-ma (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —

36.2 ku-wa-ba-shi-ri-a me-ma (8)
lexical 6·2 — P1=5, P2=7 — G-lex 2
auditory 6·2 — P1=5, P2=7 — G-aud 2 — Δ —
36.3 mu-m-fu-wa-te ka-ri-ma (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
36.4 ke-sho m-pa-te ku-ki-a (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ –P2=4
Stanza 37
37.1 hi-yo ndi-yo ya-ngu e-zi (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
37.2 na-po-ze-wa ni a-zi-zi (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
37.3 ku-wa mi-mi mu-o-mbe-zi (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
37.4 u-ma-ti ku-wa-o-mbe-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 38
38.1 m-fu-wa-te-ni ma-na-ni (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
38.2 na-mi m-tu-me a-mi-ni (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
38.3 ke-sho mwi-ngi-e phe-po-ni (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
38.4 hu-ri m-ka-e pa-mo-ya (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
Stanza 39
39.1 fu-wa-ta-ni ndi-lo pa-to (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —

39.2 m-fa-ha-mu ku-na nyo-to (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
39.3 hu-wa-ka kwa mi-vu-ku-to (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
39.4 ha-u-na ku-zi-zi-mi-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
Stanza 40
40.1 nyo-the ni-me-wa-khu-bi-ri (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
40.2 na ku-wa-kho-fi-sha na-ri (8)
lexical 6·2 — P1=5, P2=7 — G-lex 2
auditory 6·2 — P1=5, P2=7 — G-aud 2 — Δ —
40.3 mu-wa-te ku-ta-ka-ba-ri (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
40.4 m-ki-ri mo-la ja-li-a (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —

4.c. Excluded lines

Eleven lines are removed from the statistical pool. Arabized executions preserve closed syllables and geminates and thereby displace the stress positions mechanically. *Guni* of reading are verses in which the reader adds or omits a syllable that the source text does not contain, the determination being made by direct comparison with the text supplied to her. One verse is excluded because the reader, after three attempts, produced no usable execution. Genuine *guni* of the source text are retained.

1.1 *bis-mi-lla-hi a-wa-li* — Arabized execution
2.1 *a-rah-ma-ni a-zi-zi* — Arabized execution
4.1 *al-ham-du li-lla-hi* — Arabized execution
7.2 hu-mwe-nde-le-a mu-ri-sa-li — *guni* of reading
13.3 *al-ha-ji-ri ma-ka* — Arabized execution
18.4 *kamwamkuwa alia* — inaudible execution

21.2 *ka-mba a-ba al-ha-sa-ni* — Arabized execution
23.1 *ni-lo na-lo kwa-mbi-e* — *guni* of reading
25.1 *sa-sa hi-ya ta-to-ka* — *guni* of reading
31.2 *kham-si-ni kwa he-sa-bu* — Arabized execution
34.1 *bis-mi-lla-hi ra-di-di* — Arabized execution

4.d. Divergences between the source text and the execution

The reader received the Kiamu text deposited in the Zenodo repository and was instructed to read it aloud and not to sing it. Forty-seven of the 160 verses diverge from that text. The divergences fall into three classes, and only the third bears on the statistical pool.

Class 1 — systematic dialectal substitution, metrically inert. A speaker of standard Swahili from Tabora regularly renders the Kiamu cluster *dr* as *d* (*ndrio*, *ndriye*, *ndripo*, *ndrilo*, *ndrimi*, *kwendra*, *enendra*, *nendre*, *wendre*), *th* as *t* (*mthu*, *thumwa*, *zothe*, *wethu*, *thaka*, *thatoka*, *thakupa*), and *kwe* as *ke* (*yakwe*, *zakwe*). None of these substitutions alters the syllable count, and none alters the position of any stress: *ndri-o* and *ndi-o* are both disyllabic, as are *mthu* and *mtu*, *ya-kwe* and *ya-ke*.

Class 2 — lexical substitution without change of count. A small number of verses show a change of word or of grammatical person that leaves the *mizani* count intact: at 2.2 *Rabi* is read as *Nabi*, at 37.2 *Nipozewa* as *napozewa*. These are recorded but do not affect the syllabic or accentual analysis.

Class 3 — change of syllable count. Five verses, all of them excluded on that ground: at 23.1 the source *Nililo nalo khwambie* (8) is executed as *nilo nalo kwambie* (7); at 25.1 *Sasa hiyau thatoka* (8) as *sasa hiya tatoka* (7); at 31.2 *Khamusini kwa hesabu* (8) as *khamsini kwa hesabu* (7), by Arabization of the numeral; at 13.3 *Alihajiri Maka* (7) as *al-hajiri maka* (6), likewise by Arabization; and at 7.2 *Humwendrea murisali* (8) as *humwendelea murisali* (9).

Four further verses deviate from the eight-syllable count without being execution artifacts, since the deviation is already present in the source text: 5.2 *Zeneo kama mwezi* (7), 12.3 *Yalipoingia fitina* (9), 25.4 *Inshallah thakwetea* (7) and 27.3 *Thakupa na mare makamu* (9). These are *guni* of the text itself and are retained in the pool, where they count against the syllabic matrix.

The complete list of the forty-seven divergent verses follows, source form first and executed form second.

1.1 Bisimillahi awali → bis-mi-lla-hi a-wa-li
1.2 Isimu yakwe Jalali → i-si-mu ya-ke ja-la-li
1.3 Ndrio mwando wa kauli → ndi-o mwa-ndo wa ka-u-li
1.4 Wa mthu kutamkia → wa m-tu ku-ta-m-ki-ya
2.1 Arrahmani, Azizi → a-rah-ma-ni a-zi-zi
2.2 Rabi ndriye Muawazi → na-bi ndi-ye mu-a-wa-zi
4.1 Alihamdu lillahi → al-ham-du li-lla-hi
5.1 Sifa zothe njema hizi → si-fa zo-te nje-ma hi-zi
5.3 Ni zakwe yeye Azizi → ni za-ke ye-ye a-zi-zi
6.2 Ni rehema yakwe yeye → ni re-he-ma ya-ke ye-ye
6.3 Mwenye ezi Mungu ndriye → mwe-nye e-zi mu-ngu ndi-ye
7.2 Humwendrea murisali → hu-mwe-nde-le-a mu-ri-sa-li
9.1 Thumwa wethu mfadhali → tu-mwa we-tu m-fa-dha-li
9.2 Ndriye muumbwa awali → ndi-ye mu-u-mbwa a-wa-li
10.3 Ndripo khakosa maneno → ndi-po kha-ko-sa ma-ne-no
11.3 Sasa thaka kunudhumu → sa-sa ta-ka ku-nu-dhu-mu
13.3 Alihajiri Maka → al-ha-ji-ri ma-ka
13.4 Mtume wethu nabia → m-tu-me we-tu na-bi-a
14.2 Kwendra Madina bashiri → kwe-nda ma-di-na ba-shi-ri
14.4 Kwa thumwa kuwangilia → kwa tu-mwa ku-wa-ngi-li-a
16.1 Wake wothe na zijana → wa-ke wo-te na zi-ja-na
16.2 Hakwendra nao amina → ha-kwe-nda na-o a-mi-na
17.2 Ndrilo inale, wambuji → ndi-lo i-na-le wa-mbu-ji
17.4 Muhammadi kumwendrea → mu-ha-mma-di ku-mwe-nde-a
18.1 Thumwa kule Madinati → tu-mwa ku-le ma-di-na-ti
21.2 Kamba Aba lHasani → ka-mba a-ba al-ha-sa-ni
23.1 Nililo nalo, khwambie → ni-lo na-lo kwa-mbi-e
23.2 Enendra kanitwalie → e-ne-nda ka-ni-twa-li-e
23.4 Mato khiwaangalia." → ma-to ki-wa-a-nga-li-a
24.4 Thakwendra kukutwalia → ta-kwe-nda ku-ku-twa-li-a
25.1 Sasa hiyau thatoka → sa-sa hi-ya ta-to-ka
25.2 Thumwa wangu, nendre Maka → tu-mwa wa-ngu ne-nde ma-ka
25.4 Inshallah, thakwetea." → in-sha-llah ta-kwe-te-a
27.3 Thakupa na mare makamu → ta-ku-pa na ma-re ma-ka-mu
28.2 Wendre nao, mmarisi → we-nde na-o m-ma-ri-si
28.3 Kwa baba wethu upesi → kwa ba-ba we-tu u-pe-si
29.2 Wendre nao hata Maka → we-nde na-o ha-ta ma-ka
29.3 Kwa baba wethu, pulika → kwa ba-ba we-tu pu-li-ka
30.1 Na waraka hono ndrani → na wa-ra-ka ho-no nda-ni

31.2 Khamusini kwa hesabu → kham-si-ni kwa he-sa-bu
33.4 Maneno, thakokwambia → ma-ne-no ta-ko-kwa-mbi-a
34.2 Isimu yakwe Wadudi → i-si-mu ya-ke wa-du-di
34.3 Ndrimi thumwa Muhammadi → ndi-mi tu-mwa mu-ha-mma-di
37.1 Hiyo ndriyo yangu ezi → hi-yo ndi-yo ya-ngu e-zi
37.2 Nipozewa ni Azizi → na-po-ze-wa ni a-zi-zi
39.1 Fuwatani, ndrilo pato → fu-wa-ta-ni ndi-lo pa-to
40.1 Nyuthe nimewakhubiri → nyo-the ni-me-wa-khu-bi-ri

4.e. Summary

The pool comprises **149 verses** of the 160 mapped. The syllabic constant N = 8 holds in 144 of them, that is 96.6 percent; the five remaining verses are *guni* of the source text, retained in the pool.

Distribution of G-lex, derived from the printed lines: G = 1 in 1 verse (0.7 percent), G = 2 in 82 (55.0), G = 3 in 57 (38.3), G = 4 in 9 (6.0).

Distribution of G-aud, established by ear in the formal reading: G = 1 in 2 verses (1.3 percent), G = 2 in 89 (59.7), G = 3 in 51 (34.2), G = 4 in 7 (4.7).

The *Grenzschema* appears as a lexical sequence 4·4 in **17 verses (11.4 percent)**, and is realized as P_1=3, P_2=7 with G-aud = 2 in **23 verses (15.4 percent)**. A tetrasyllabic nucleus of any kind occurs in 43 verses (28.9 percent), and the nuclei of a verse are all of equal volume — the condition under which a repeating foot could exist — in 27 verses (18.1 percent).

The transitions from G-lex to G-aud comprise **16 unrealized nuclei and 5 added prominences**. The unrealized nuclei fall on position 1 (7 occurrences), position 4 (5) and position 5 (4); no nucleus on position 7 is ever unrealized. Fourteen verses carry a positional index marked as uncertain.

The five added prominences are the only verses in which the voice produces a nucleus the lexicon does not predict:

5.4 — a nucleus on position 1
7.3 — a nucleus on position 3
9.4 — a nucleus on position 3
12.2 — a nucleus on position 3
31.3 — a nucleus on position 1

4.f. Intra-annotator control

The auditory annotation is the work of a single listener. A second pass over the corpus was carried out at a remove from the first, and its results are reported here in full.

Of the 149 verses in the pool, the auditory annotation is unchanged in 130 and differs in 19, giving an agreement rate of **87.2 percent**. The direction of the disagreements is as instructive as their number: G-aud rises in 11 verses and falls in 7. The second pass is therefore not systematically more conservative than the first, and no drift in either direction is apparent.

Eleven further verses were altered on grounds that are not perceptual and are counted separately, since including them would confuse editorial decisions with listening. Six concern exclusion or retention, following comparison with the source text supplied to the reader. Four concern the terminal stress of *guni* verses, where the penultimate of the final word falls on position 6 or 8 rather than 7. One follows a philological arbitration on the enclitic locative of *ulimwenguni*.

One property of the notation must be declared, because it bears directly on the figure above. The second pass was conducted with the confidence mark available, and the first was not. A mark that records hesitation rather than resolving it makes it easier to retain a weak nucleus, and equally easier to set one aside, since neither decision is final. Part of the observed movement is therefore attributable to the instrument rather than to the ear, and the fourteen verses that now carry an uncertain positional index are the visible trace of that. An agreement rate obtained under two different notations measures the reproducibility of a judgment made under changing conditions, and it should be read as an upper estimate of the instability rather than as a measure of the listener's inconsistency.

None of this substitutes for a second annotator, and no claim of inter-annotator reliability is made. The raw audio is deposited so that any reader may supply one.

5. Statistical Analysis and Falsification of the Accentual Model

All figures are computed on the pool of 149 verses established in section 4. G-lex is derived from the printed lines by the penultimate rule and requires no recording; G-aud is the realization established by ear. The prose baseline invoked below is that of section 9.7.

5.a. Executing Conditions 1 and 2: the syllabic matrix and the rhyme

Condition 1, the syllabic constant. The matrix requires S = 8 for every line. The constant holds in **144 of the 149 retained verses, that is 96.6 percent**. The five exceptions are *guni* of the source text: 5.2 and 25.4 at seven syllables, 12.3 and 27.3 at nine, and 27.4 at seven. As set out in the condition, the treatment of these was fixed before the count. They are deviations of the text and not of the execution, they are retained in the pool, and they count against the matrix.

The condition is satisfied on its stated criterion. Deviation is rare, and the deviant lines are *guni* in the endogenous sense — defective in measure, and identified as such by the tradition that defines the term (Abedi, 1954, p. 16). Eleven further verses were removed from the pool for reasons of execution and not of composition: eight Arabized executions, in which the reader restores closed syllables and geminates that the assimilated Swahili form does not carry, two *guni* of reading, and one verse for which no usable execution was obtained. Every one of these determinations is made by direct comparison with the text supplied to the reader, and the comparison is published in full in section 4.

Condition 2, the rhyme variables. The matrix requires a local rhyme *a* binding the first three lines of each stanza and a global rhyme *X*, the *bahari*, closing every fourth line across the whole poem.

The local rhyme holds in **39 of the 40 stanzas**. The single exception is Stanza 10, where the three lines close on *kwenu*, *mno* and *maneno*: the rhyming syllable is *-nu* in the first and *-no* in the two others. The divergence is present in the source manuscript and is not an artifact of the reading.

The global rhyme. The matrix and algebraic equation of section 2 leave the global constant unresolved, written *X*, and the two forms are tested in the order fixed by the symbolic alphabet.

X^{s}, the syllabic global rhyme, requires that every fourth line close on the same full consonant-vowel sequence. It is contradicted by the corpus. Thirty-four of the forty anchored lines close on the isolated vocalic syllable *-a*, but six do not: five close on *-ya* (1.4, 5.4, 10.4, 32.4, 38.4) and one on *-wa* (30.4). The onset varies, and the syllabic identity that X^{s} demands does not hold.

X^{v}, the vocalic global rhyme, requires the identity of the terminal vowel alone, the onset being free to vary. It holds in **40 stanzas of**

40. Every fourth line closes on the vowel /a/, whether realized as an isolated vocalic syllable or integrated within a varying consonant-vowel onset, without exception across the whole of the mapped corpus.

The disjunction is therefore resolved as X^{v}, and the algebraic equation of this corpus reads $\boldsymbol{Utenzi} = \mathbf{3}(\mathbf{8}_{a}\ //) \cdot \mathbf{8}_{X^{v}}\ //$. The resolution is a result of the confrontation and not a choice made to accommodate it: X^{s} was tested, the six lines that contradict it are named above, and X^{v} is what the text imposes in its place.

Conditions 1 and 2 are therefore satisfied, with one imperfect local rhyme, five *guni* declared, and a vocalic global rhyme. It should be stated plainly what this establishes and what it does not. These conditions verify the exhaustiveness of a description derived from the endogenous theorists and applied to a text they did not analyze; they do not falsify a rival hypothesis. The falsification is carried by Conditions 5 to 8, which follow.

5.b. Executing Condition 5: the bipartite accentual axiom (G = 2)

Condition 5 asks whether the eight-syllable line carries exactly two accentual groups, as Axiom 1 of section 2.5.1 requires. The question is put twice, to the lexicon and to the reading.

The lexicon of the line yields exactly two accentual groups in **82 verses of 149, that is 55.0 percent**; one group in 1 verse, three in 57, four in 9. The formal reading realizes exactly two prominences in **89 verses, that is 59.7 percent**; one in 2, three in 51, four in 7.

The result must be stated without embellishment, because it does not go entirely against the accentual model. The bipartite count is the majority outcome on both measures, and any account of this corpus which denied that two accents are frequent would be contradicted by its own data. No such claim is made here.

What the figures establish is that the count is not a law. As an axiom, G = 2 admits no exception, and this corpus supplies 67 counter-examples on the lexical measure and 60 on the auditory one. As a statistical tendency, it holds in a little over half the lines. A generative rule violated in two lines out of five is not a rule but a tendency, and a tendency requires an account of its exceptions that the accentual school has never supplied.

The transition disposes of the objection that the count is an artifact of the analyst's ear. Of the 82 verses whose lexicon already yields two groups, 78 are realized with two: where the text

offers the bipartite configuration, the reading takes it. Of the 66 verses whose lexicon yields one, three or four groups, **only 10 are brought to two in the reading**, that is 15.2 percent; 56 are not. The voice does not drive the line toward a binary pulse. It very largely preserves the accentual structure the lexicon supplies, and the rise from 55.0 to 59.7 percent is produced by ten verses.

The prose baseline settles what this movement means. In the control corpus of section 9.7, read by the same speaker under the same instruction, the reading departs from the lexical structure in 12.5 percent of lines; in this corpus it departs in 12.1 percent. A reader alters the accentual structure of one line in eight, at the same rate, whether she is reading a nineteenth-century *utendi* or a sentence about the price of fish. The reduction of G-lex to G-aud is therefore not a poetic phenomenon at all, and it cannot be what a meter does to a line.

This bears directly on the mechanism the accentual model requires. In the 66 verses where the lexicon yields other than two groups, Axiom 1 can be satisfied only if the performer suppresses or adds a nucleus — that is, only if the performance departs from the accentual structure of the words. This is the operation observed in *kwa mámbo bora adhímu*, discussed in section 3.3.3, where the bipartite layout appears solely because the natural stress of *bora* has been swallowed. The reading examined here performs that operation in 10 verses out of 66 and declines to perform it in 56. The bipartite configuration is not something this text offers; in nearly half the corpus it is something a performance would have to impose.

Deletions and additions. The transitions comprise 16 unrealized nuclei and 5 added prominences. The unrealized nuclei fall on position 1 (seven occurrences), position 4 (five) and position 5 (four). **No nucleus on position 7 is ever unrealized.** The five additions are the only verses in which the voice produces a nucleus the lexicon does not predict: at 5.4 and 31.3 on position 1, and at 7.3, 9.4 and 12.2 on position 3. Those three verses on position 3, out of 149, are the whole of the positive evidence this corpus offers for a beat that is not derivable from the lexicon. Fourteen verses carry a positional index marked as uncertain.

5.c. Executing Condition 6: the *Grenzschema* as a cognitive prototype

Condition 6 concerns not the number of the groups but their volumes and positions. The *Grenzschema* is the ideal matrix of section

2.5.1: $|A_1| = 4$, $|A_2| = 4$, with $P_1 = 3$ and $P_2 = 7$. In the notation of nucleus volumes it is the single sequence 4·4.

The lexicon of the *Utendi wa Haudaji* produces that sequence in **17 verses of 149, that is 11.4 percent**. It is realized as $P_1 = 3$, $P_2 = 7$ in **23 verses, that is 15.4 percent**. The prototype fails to reach one line in six on either measure.

The prose baseline removes the prototype's claim to be metrical. In the control corpus, the eight lines whose lexicon yields the sequence 4·4 are realized with $P_1 = 3$ and $P_2 = 7$ in all eight cases. The configuration arises automatically, in ordinary prose, whenever the lexicon supplies two words of four syllables. It requires no metrical intention and no poetic context. What Condition 6 tests, then, is not whether the configuration exists — it plainly does, in prose as in verse — but whether the poem produces it more often than the lexicon alone would. At 11.4 percent lexically, it does not.

One divergence is recorded rather than passed over. The poem moves from 17 lexical sequences 4·4 to 23 realizations of $P_1 = 3$, $P_2 = 7$: seven verses converge on the configuration without the lexicon supplying it, arriving from five different lexical sequences — 2·2·4, 4·2·2, 2·2·2·2, 5·3, and one eight-syllable word. The prose corpus shows no such convergence: its eight realizations are exactly its eight lexical 4·4 lines. The numbers are small, no single mechanism accounts for the seven verses, and the observation is reported here as the only measurement in this study that distinguishes the poem from prose.

The decisive figure comes from within the accentual model's own successes. Among the 89 verses that satisfy Axiom 1, the two prominences are distributed as follows: on positions 2 and 7 in 30 verses, on 3 and 7 in 23, on 4 and 7 in 16, on 1 and 7 in 15, on 5 and 7 in 2, on 2 and 6 in 2, on 5 and 8 in 1. **Fewer than a quarter of the verses that vindicate the accentual school on the number vindicate it on the position**, and the dominant configuration is not $P_1 = 3$ but $P_1 = 2$. In the prose corpus the dispersion is the same, and there too no first position is favored. A prototype realized in one case out of four among the cases most favorable to it, and no more often than in prose, is not a cognitive attractor.

5.d. Executing Condition 7: the absence of periodicity

Condition 7 is where the accentual model fails without remainder, and it fails on ground that requires no recording at all.

Intra lines. A metrical foot is defined by recurrence: it is a unit that repeats. A line therefore exhibits an internal periodicity only if its accentual nuclei are of equal volume, that is, if |A1| = |A2| = … = |A*G*|. The sequences 4·4 and 2·2·2·2 satisfy this; 3·5 and 2·2·4 do not, since they juxtapose units of different sizes, which no metrical system would call feet.

In the *Utendi wa Haudaji*, the nuclei of a verse are all of equal volume in **27 verses of 149, that is 18.1 percent**. In the remaining 122 there is no repeating unit whatever, and consequently no foot.

The 82 verses whose lexicon yields exactly two groups distribute as follows, and every figure here is computed from the printed text by the penultimate rule, depending on no recording, no reader and no ear: the sequence 3·5 in 27 verses, with the stresses on positions 2 and 7 and an interval of five syllables; 5·3 in 18, on positions 4 and 7 with an interval of three; 2·6 in 15, on 1 and 7 with an interval of six; 4·4 in 17, on 3 and 7 with an interval of four; 6·2 in 2, on 5 and 7 with an interval of two; and three verses in 3·4 and 6·3 among the *guni*.

The first stressed syllable falls on position 1, 2, 3, 4 or 5, and the interval separating the two stresses varies from two to six syllables. There is no recurrent unit and no constant interval. Two beats at a variable distance do not constitute a rhythm; they constitute two beats.

This result is independent of everything that could reasonably be disputed in this study. Even if Axiom 1 were granted in full, even if every verse of the corpus were conceded to carry exactly two accentual groups, no periodic structure would follow and no trochaic foot could be recovered. The accentual school does not merely fail to predict the number of the beats; it fails to predict their placement in the very verses where the number is correct.

Across lines. The second scale of Condition 7 asks whether the poet sequences these configurations, and its criterion is independence: the observed frequency of identical sequences of nucleus volumes in consecutive lines is compared with the frequency expected if each line's configuration were drawn independently, given the distribution of configurations in the corpus. Across the 148 consecutive pairs of the mapped corpus, two successive lines carry the same sequence of nucleus volumes in **17 pairs, that is 11.5 percent**. Independence predicts 15.3 pairs, that is 10.4 percent. The observed frequency exceeds the expected one by fewer than two pairs in a hundred and forty-eight. Measured on the number of groups alone, disregarding their volumes, the figures are 72 pairs observed against

67 expected. The same sequence recurs across three consecutive lines twice in the whole corpus. Within the stanza, where a compositional effect would be most expected, 15 pairs of 109 are identical, that is 13.8 percent, against 10.4 percent for the corpus at large. There is no sequencing. The succession of accentual profiles from one line to the next is what independent draws from the corpus distribution would produce, which is to say that it is governed by the morphological size of the words the poet selected to fill the syllabic count, and by nothing else. A generative metrical law dictates the words; here the words dictate the accents.

5.e. Executing Condition 8: the tetrasyllabic foot

Condition 8 tests Minerba's derivation of the Swahili line from an imported tetrasyllabic template against the metrical boundaries the corpus actually imposes. It must be said at once that the *utenzi* cannot falsify it. The line counts eight *mizani*, which is a multiple of four, so a partition into two tetrasyllabic units is arithmetically available in every regular verse of this corpus. The condition is decisive only where a segment cannot accommodate the template at all, which is the case of the rupture and hyper-fractionated *wimbo* matrices examined in Case Studies III and V, and it is tested there. What this case study can establish is narrower: whether the template, being available, is actually used.

It is not. A tetrasyllabic nucleus of any kind occurs in **43 verses of 149, that is 28.9 percent**. Of those 43, only **17** contain two, which is to say only 17 are partitioned as 4·4 and could be described as two tetrasyllabic feet. The remaining 26 contain a single tetrasyllabic nucleus flanked by units of two or three syllables — sequences such as 2·2·4, 4·2·2 or 2·4·2 — in which the four-syllable group is not a repeating unit but an accident of one long word. In 106 verses, that is 71.1 percent, no tetrasyllabic nucleus occurs at all.

The template is therefore available in every line and taken up in one line in nine. A metrical unit that the language permits everywhere and that the composer employs in eleven percent of his verses is not the generative principle of the meter; it is one of the shapes the lexicon happens to produce. The prose control corpus of section 9.7 makes the same point from the other side: there too the sequence 4·4 appears wherever two four-syllable words fall together, and there too it appears in no other way.

5.f. Analytical conclusion: the terminal anchor

One constant survives every measurement in this case study, and it is the one the accentual school has consistently mistaken for evidence. Across the 149 lines a nucleus falls on position 7 in every regular line, and it is never unrealized in the reading: not once, in a corpus where nuclei on positions 1, 4 and 5 are lost sixteen times between them.

The invariance is real. It carries no metrical information whatever.

It follows mechanically from the penultimate rule applied to the last word of the line. Classical Swahili poetics prohibits enjambment, so the line ends on a word boundary; that word carries its lexical stress on its penultimate syllable; and in an eight-syllable line the penultimate of the final word is position 7. Stated lexically rather than positionally — the terminal stress is the penultimate of the last word — the invariant holds without a single exception in this corpus, including in the *guni*, where it falls on position 6 or 8 according to the length of the line. In the prose control corpus the same invariant holds: the second prominence of every two-group line falls on position 7, in sentences that have no meter at all.

Axiom 2 of section 2.5.1 concedes as much, since Vierke derives $P_2 = 7$ from the penultimate stress of the phonological word terminating the line. The axiom therefore states a fact of Swahili morphology and presents it as a metrical law.

To take this by-product of word structure, observe that it recurs at the end of every line, and infer from it a trochaic nucleus toward which the verse is built, is to draw the target around the shot after it has been fired. The regularity is genuine; its cause is phonological, not metrical; and it obtains identically in Swahili prose, as section 9.7 establishes on a corpus written for the purpose. What remains of the accentual model, once this single invariant is subtracted, is a count that holds in a little over half the lines and a placement that holds in fewer than a quarter of those.

6. Instrumental Verification of the Hostile Sample

Instrumental analysis is applied to a sample of verses and to no more than that, as section 9.2.3 sets out. The exhaustive verification of this corpus is auditory, and no statistic in this case study rests on the measurements reported below. What the sample establishes is that the categories counted in section 5 have measurable physical correlates, and where they do not, that too is reported.

6.a. Composition of the sample

The composition was fixed before any measurement was made, and it follows the six requirements stated in section 9.2.3, one verse for each, together with two verses drawn at random. It deliberately includes the ground most favorable to the accentual model.

A verse conforming exactly to the *Grenzschema*. Stanza 37, line 2, *napozewa ni azizi*: lexical sequence 4·4, stresses on positions 3 and 7. This is the accentual model's ideal matrix, and it is measured first so that the model is tested where it is strongest.

An asymmetrical verse yielding G-aud = 2. Stanza 21, line 3, *mno nimewatamani*: sequence 2·6, stresses on positions 1 and 7. The count agrees with the bipartite axiom and the positions do not, with the two prominences six syllables apart.

A verse yielding G-aud = 3. Stanza 37, line 3, *kuwa mimi muombezi*: sequence 2·2·4, stresses on positions 1, 3 and 7.

A verse yielding G-aud = 4. Stanza 32, line 4, *kula yambo moyamoya*: sequence 2·2·2·2, stresses on positions 1, 3, 5 and 7.

A verse carrying an addition. Stanza 9, line 4, *yasiyomtindikia*: a single phonological word of eight syllables, whose lexicon yields one nucleus on position 7. The reading realizes a second prominence on position 3, which the lexicon does not predict. This is one of the five additions in the whole corpus and one of only three that fall on position 3; it is the strongest positive evidence the corpus offers to the accentual model, and it is measured for that reason.

A verse carrying a nucleus judged weak. Stanza 22, line 1, *usiku silali tena*: sequence 3·3·2, with the nucleus on position 5 marked as uncertain in the auditory annotation. The instrument reports on a decision the ear did not settle, without revising it.

Two verses drawn at random. Stanza 5, line 3, *ni zake yeye azizi*, and stanza 20, line 4, *siche kunikarifia*. The draw was made once, on the population of 149 retained verses ordered by stanza and then by line, with a seed derived from the ISBN of this book as described in section 9.2.3. Verses already belonging to the reasoned sample were skipped and the next in the generated sequence taken. The two verses obtained are analyzed and published whatever their profile, and no second draw was made.

6.b. What is measured and what is reported

For each verse of the sample, three parameters are measured syllable by syllable: fundamental frequency in semitones relative to

the speaker's median and corrected for the declination internal to the line, syllable duration, and mean intensity. The analysis range is derived from the speaker's own distribution rather than set by hand, and the conditioning applied to the F0 track is stated in the header of every script.

Each figure carries three readings of the same verse. The lexical reading is derived from the printed line by the penultimate rule. The auditory reading is the annotation of section 4. The instrumental reading is established without reference to either: a syllable counts as instrumentally prominent when its detrended F0 mean is a local maximum across the syllables of the line and lies above the declination. That criterion is free to confirm a nucleus the lexicon predicts, to leave one unconfirmed, or to designate a syllable that neither the lexicon nor the ear proposes.

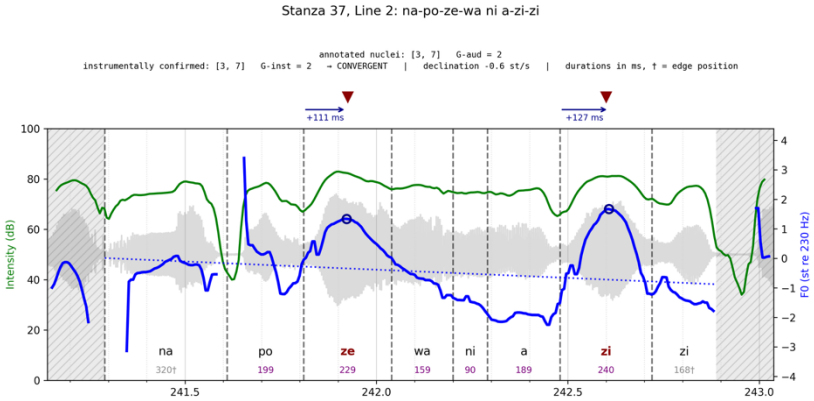
Where the three readings diverge, the divergence is stated in the caption, position by position, and the auditory annotation is not revised. Fundamental frequency is one parameter of three and is not decisive on its own; the durations and intensities are printed beneath every figure whether or not they agree with it.

6.c. Reproduction

One script corresponds to each verse of the sample. Each carries in its header the accentual analysis of the verse, the boundaries of the line in the recording and the onset of every syllable, and each produces its figure and its caption from the audio alone. The scripts and the recording are deposited as described in section 9.2.7.

6.d. Figures

Stanza 37, Line 2, Napozewa ni Azizi

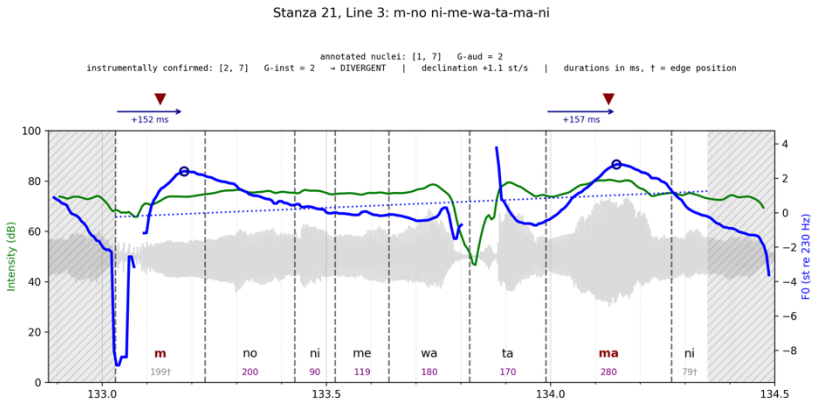



```
Utendi wa Haudaji, Stanza 37, Line 2: "na-po-ze-wa ni a-zi-zi"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4·4, P₁=3, P₂=7, G-aud = 2, delta —.
Instrumental reading: 4·4, P₁=3, P₂=7, G-inst = 2, delta — (against
the lexical reading).
Speaker median 230 Hz; declination −0.56 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 na     320 ms†  F0 -0.35 st  peak +191 ms  intensity  71.8 dB
  2 po     199 ms   F0 +0.10 st  peak + 44 ms  intensity  67.6 dB
  3 ze     229 ms   F0 +1.07 st  peak +111 ms  intensity  78.3 dB
  4 wa     159 ms   F0 -0.26 st  peak +  4 ms  intensity  76.3 dB
  5 ni      90 ms   F0 -0.98 st  peak + 39 ms  intensity  74.5 dB
  6 a      189 ms   F0 -1.37 st  peak +189 ms  intensity  73.3 dB
  7 zi     240 ms   F0 +1.18 st  peak +127 ms  intensity  76.8 dB
  8 zi     168 ms†  F0 -0.45 st  peak + 32 ms  intensity  72.5 dB
```

STANZA 21, LINE 3, MNO NIMEWATAMANI

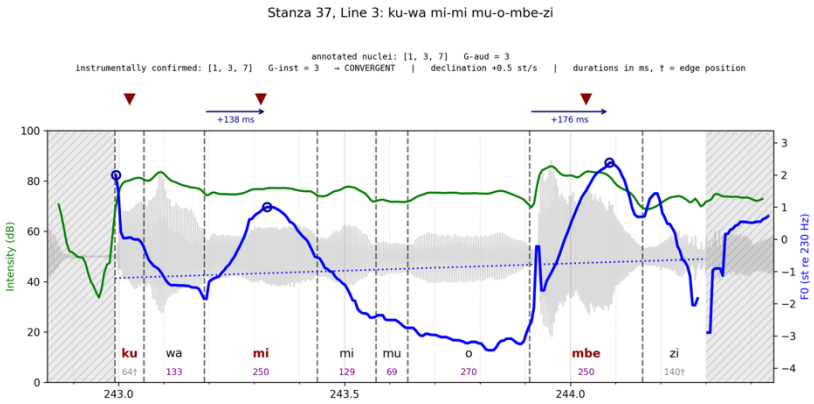



```
Utendi wa Haudaji, Stanza 21, Line 3: "m-no ni-me-wa-ta-ma-ni"
Lexical reading: 2·6, P₁=1, P₂=7, G-lex = 2.
Auditory reading: 2·6, P₁=1, P₂=7, G-aud = 2, delta —.
Instrumental reading: 3·5, P₁=2, P₂=7, G-inst = 2, delta −P₁=1, +P₁=2
(against the lexical reading).
Speaker median 230 Hz; declination +1.13 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [1] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 m      199 ms†  F0 -0.25 st  peak +152 ms  intensity  71.3 dB
  2 no     200 ms   F0 +1.10 st  peak +  2 ms  intensity  75.7 dB
  3 ni      90 ms   F0 -0.02 st  peak +  3 ms  intensity  75.6 dB
  4 me     119 ms   F0 -0.46 st  peak +  2 ms  intensity  75.0 dB
  5 wa     180 ms   F0 -0.91 st  peak +122 ms  intensity  72.9 dB
  6 ta     170 ms   F0 -0.75 st  peak + 59 ms  intensity  70.0 dB
  7 ma     280 ms   F0 +0.64 st  peak +157 ms  intensity  77.3 dB
  8 ni      79 ms†  F0 -0.90 st  peak +  5 ms  intensity  74.5 dB
```

STANZA 37, LINE 3, KUWA MIMI MUOMBEZI

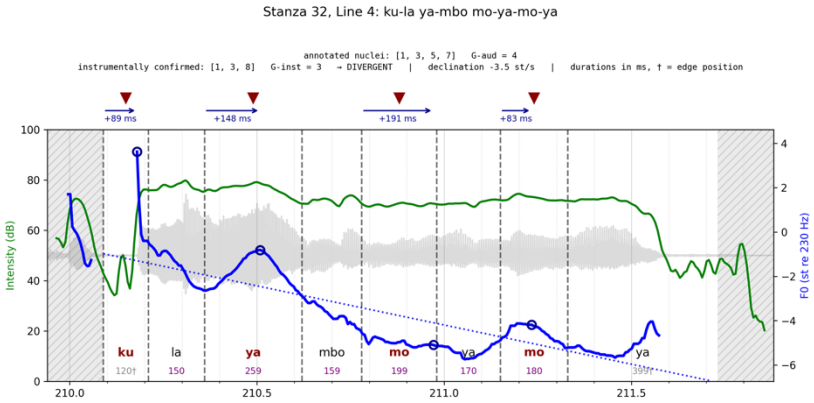



```
Utendi wa Haudaji, Stanza 37, Line 3: "ku-wa mi-mi mu-o-mbe-zi"
Lexical reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-lex = 3.
Auditory reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-aud = 3, delta —.
Instrumental reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-inst = 3, delta —
(against the lexical reading).
Speaker median 230 Hz; declination +0.45 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 ku      64 ms†  F0 +1.49 st  peak +  1 ms  intensity  79.5 dB
  2 wa     133 ms   F0 -0.12 st  peak +  4 ms  intensity  79.5 dB
  3 mi     250 ms   F0 +1.08 st  peak +138 ms  intensity  75.9 dB
  4 mi     129 ms   F0 -0.65 st  peak +  5 ms  intensity  75.3 dB
  5 mu      69 ms   F0 -1.59 st  peak +  3 ms  intensity  71.9 dB
  6 o      270 ms   F0 -2.25 st  peak +  0 ms  intensity  74.6 dB
  7 mbe    250 ms   F0 +1.45 st  peak +176 ms  intensity  80.3 dB
  8 zi     140 ms†  F0 +0.72 st  peak + 26 ms  intensity  71.6 dB
```

STANZA 32, LINE 4, KULA YAMBO MOYA MOYA



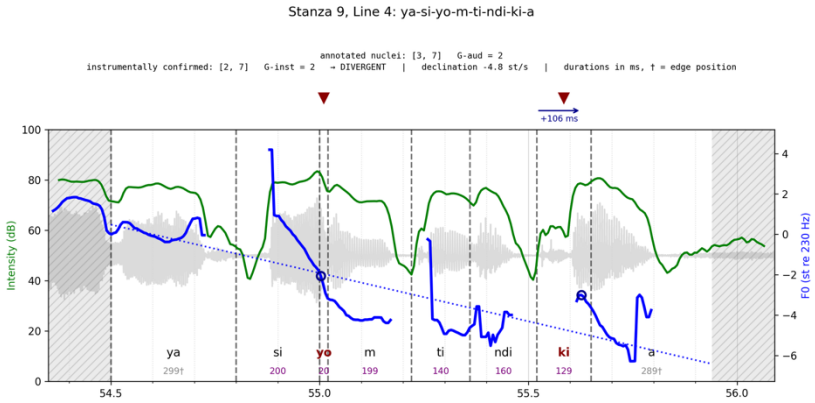
Utendi wa Haudaji, Stanza 32, Line 4: "ku-la ya-mbo mo-ya-mo-ya"
Lexical reading: 2·2·2·2, P_1=1, P_2=3, P_3=5, P_4=7, G-lex = 4.
Auditory reading: 2·2·2·2, P_1=1, P_2=3, P_3=5, P_4=7, G-aud = 4, delta —
Instrumental reading: 2·2·4, P_1=1, P_2=3, P_3=8, G-inst = 3, delta $-P_3$=5, $-P_4$=7, $+P_3$=8 (against the lexical reading).
Speaker median 230 Hz; declination −3.53 st/s.

The instrumental reading diverges from the auditory reading: position(s) [5, 7] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [8] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	ku	120 ms†	F0 +1.95 st	peak + 89 ms	intensity	51.4 dB
2	la	150 ms	F0 +0.12 st	peak + 3 ms	intensity	77.3 dB
3	ya	259 ms	F0 +0.63 st	peak +148 ms	intensity	76.2 dB
4	mbo	159 ms	F0 -0.54 st	peak + 0 ms	intensity	71.8 dB
5	mo	199 ms	F0 -1.21 st	peak +191 ms	intensity	70.1 dB
6	ya	170 ms	F0 -0.93 st	peak +169 ms	intensity	70.8 dB
7	mo	180 ms	F0 +0.52 st	peak + 83 ms	intensity	72.7 dB
8	ya	399 ms†	F0 +0.62 st	peak +226 ms	intensity	60.8 dB

STANZA 9, LINE 4, YASIYOMTINDIKIA



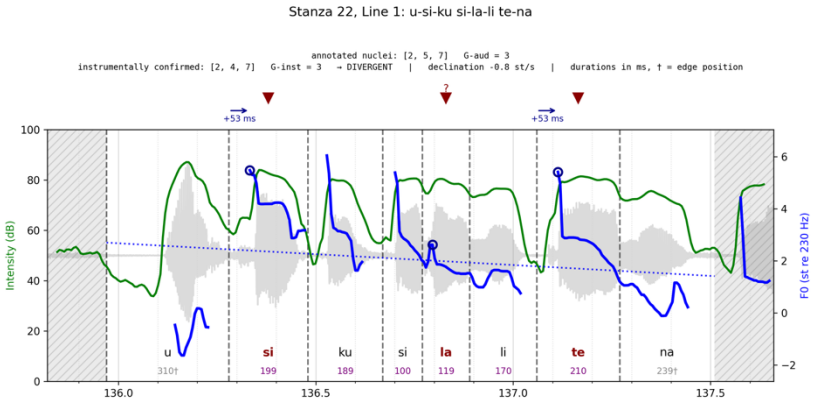
Utendi wa Haudaji, Stanza 9, Line 4: "ya-si-yo-m-ti-ndi-ki-a"
Lexical reading: 8, P_1=7, G-lex = 1.
Auditory reading: 4·4, P_1=3, P_2=7, G-aud = 2, delta +P_1=3.
Instrumental reading: 3·5, P_1=2, P_2=7, G-inst = 2, delta +P_1=2 (against the lexical reading).
Speaker median 230 Hz; declination −4.79 st/s.

The instrumental reading diverges from the auditory reading: position(s) [3] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	ya	299 ms†	F0 +0.19 st	peak +207 ms	intensity	71.4 dB
2	si	200 ms	F0 +1.75 st	peak + 85 ms	intensity	68.3 dB
3	yo	20 ms	F0 -0.68 st	peak + 2 ms	intensity	78.4 dB
4	m	199 ms	F0 -1.63 st	peak + 10 ms	intensity	67.2 dB
5	ti	140 ms	F0 -0.84 st	peak + 39 ms	intensity	68.6 dB
6	ndi	160 ms	F0 -0.68 st	peak + 88 ms	intensity	63.4 dB
7	ki	129 ms	F0 +1.69 st	peak +106 ms	intensity	66.3 dB
8	a	289 ms†	F0 +0.65 st	peak +116 ms	intensity	61.5 dB

STANZA 22, LINE 1, USIKU SILALI TENA



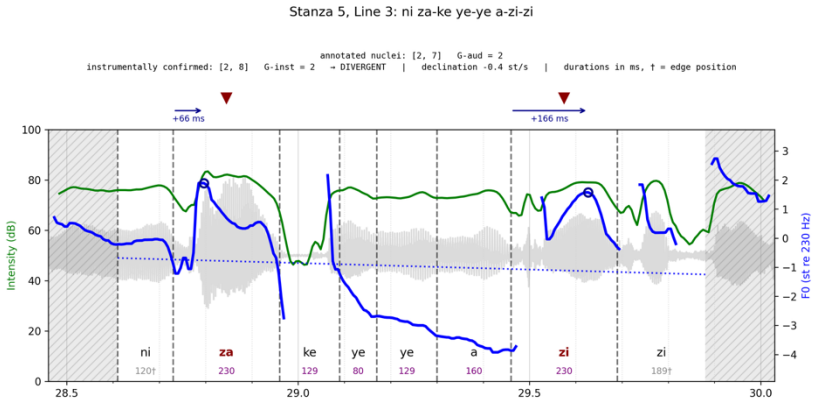
Utendi wa Haudaji, Stanza 22, Line 1: "u-si-ku si-la-li te-na"
Lexical reading: 3·3·2, P_1=2, P_2=5, P_3=7, G-lex = 3.
Auditory reading: 3·3·2, P_1=2, P_2=5?, P_3=7, G-aud = 3, delta –.
Instrumental reading: 3·2·3, P_1=2, P_2=4, P_3=7, G-inst = 3, delta –P_2=5, +P_2=4 (against the lexical reading).
Speaker median 230 Hz; declination –0.84 st/s.

The instrumental reading diverges from the auditory reading: position(s) [5] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [4] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	u	310 ms†	F0 -3.20 st	peak +234 ms	intensity	59.1 dB
2	si	199 ms	F0 +1.67 st	peak + 53 ms	intensity	73.0 dB
3	ku	189 ms	F0 +0.86 st	peak + 48 ms	intensity	66.0 dB
4	si	100 ms	F0 +0.93 st	peak + 30 ms	intensity	73.4 dB
5	la	119 ms	F0 -0.18 st	peak + 25 ms	intensity	78.0 dB
6	li	170 ms	F0 -0.57 st	peak + 72 ms	intensity	68.6 dB
7	te	210 ms	F0 +0.96 st	peak + 53 ms	intensity	74.0 dB
8	na	239 ms†	F0 -0.94 st	peak +144 ms	intensity	67.1 dB

STANZA 5, LINE 3, NI ZAKE YEYE AZIZI



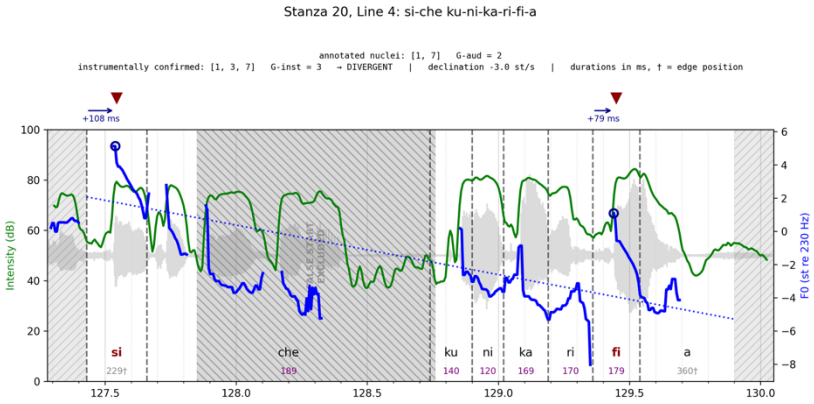
Utendi wa Haudaji, Stanza 5, Line 3: "ni za-ke ye-ye a-zi-zi"
Lexical reading: 3·2·3, P_1=2, P_2=4, P_3=7, G-lex = 3.
Auditory reading: 3·5, P_1=2, P_2=7, G-aud = 2, delta $-P_2$=4.
Instrumental reading: 3·5, P_1=2, P_2=8, G-inst = 2, delta $-P_2$=4, $-P_3$=7, $+P_2$=8 (against the lexical reading).
Speaker median 230 Hz; declination −0.45 st/s.

The instrumental reading diverges from the auditory reading: position(s) [7] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [8] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	ni	120 ms†	F0 +0.54 st	peak + 85 ms	intensity	76.5 dB
2	za	230 ms	F0 +1.14 st	peak + 66 ms	intensity	77.1 dB
3	ke	129 ms	F0 +0.21 st	peak +103 ms	intensity	56.5 dB
4	ye	80 ms	F0 -1.22 st	peak + 1 ms	intensity	75.2 dB
5	ye	129 ms	F0 -1.96 st	peak + 5 ms	intensity	73.5 dB
6	a	160 ms	F0 -2.60 st	peak + 3 ms	intensity	73.4 dB
7	zi	230 ms	F0 +1.64 st	peak +166 ms	intensity	74.3 dB
8	zi	189 ms†	F0 +1.66 st	peak + 53 ms	intensity	66.8 dB

STANZA 20, LINE 4, SICHE KUNIKARIFIA




```
Utendi wa Haudaji, Stanza 20, Line 4: "si-che ku-ni-ka-ri-fi-a"
Lexical reading: 2·6, P₁=1, P₂=7, G-lex = 2.
Auditory reading: 2·6, P₁=1, P₂=7, G-aud = 2, delta −.
Instrumental reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-inst = 3, delta +P₂=3
(against the lexical reading).
Speaker median 230 Hz; declination −2.98 st/s.
```

Measurement exclusion: the acoustic interval [(127.85, 128.76)] was omitted due to a false start. Because this line was designated by the random seed of the hostile sample, it must be retained; the aborted attempt is excluded from all measurements to preserve structural integrity.

The instrumental reading diverges from the auditory reading: position(s) [3] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 si     229 ms†  F0 +1.30 st  peak +108 ms  intensity  67.8 dB
  2 che    189 ms   F0 -1.19 st  peak + 74 ms  intensity  65.1 dB
  3 ku     140 ms   F0 +0.23 st  peak +114 ms  intensity  58.9 dB
  4 ni     120 ms   F0 -0.93 st  peak + 15 ms  intensity  78.8 dB
  5 ka     169 ms   F0 -0.61 st  peak + 68 ms  intensity  72.8 dB
  6 ri     170 ms   F0 -1.11 st  peak +116 ms  intensity  68.9 dB
  7 fi     179 ms   F0 +2.83 st  peak + 79 ms  intensity  72.3 dB
  8 a      360 ms†  F0 +0.29 st  peak +134 ms  intensity  58.9 dB
```

6.f. The concordance of the instrumental reading

The protocol of section 9.2.3 requires that this step report a concordance rate together with its failures. Of the eight verses measured, the instrumental reading coincides exactly with the auditory annotation in **two**, at 37.2 and 37.3, and departs from it in **six**. The rate is therefore two of eight, and the departures are set out below rather than summarized away.

The second of the two concordances deserves a sentence of its own. At 37.3, *kuwa mimi muombezi*, lexicon, ear and instrument converge on three accentual groups, 2·2·4, at positions 1, 3 and 7, the same configuration as at 12.4 of the *Rasi 'lGhuli*: an eight-syllable line carrying the configuration the bipartite axiom declares impossible, confirmed by every procedure this study employs.

The disagreement is not evenly distributed between the two things an accentual reading asserts. On the **number** of nuclei, instrument and ear agree in six verses of eight, disagreeing only at 32.4, where the instrument finds three where the ear finds four, and at 20.4, where it finds three where the ear finds two. On the **placement** of those nuclei they agree in two verses only. The instrument is thus a fair witness to how many prominences a line carries and a poor one to where they fall, which is precisely the distinction Condition 5 and Condition 6 keep apart.

The six departures fall into three kinds, and each has an identifiable cause.

Three are displacements of a single syllable, and they do not all run the same way. At 21.3 the instrument places the first nucleus on position 2 where the ear places it on 1: the ear's syllable is the syllabic nasal *m*, which carries no vowel, and the pitch movement is realized on the vowel that follows. At 9.4 and 22.1 the instrument moves the nucleus one syllable earlier, to position 2 where the ear places it on 3, and to position 4 where the ear places it on 5. At 9.4 the ear's syllable, *yo*, lasts 20 milliseconds and cannot carry a pitch movement of its own; at 22.1 it is both the longer and the louder of the two (119 milliseconds against 100, 78.0 dB against 73.4). Read syllable by syllable, the contour follows pitch alone; the ear weighs pitch, length and loudness together.

Two are terminal displacements. At 32.4 and at 5.3 the instrument places the final nucleus on position 8 rather than 7. Both verses close a stanza or approach its close, where the final lengthening and the boundary tone raise the last syllable above the seventh. The lexical

stress remains on the penultimate of the final word in both cases; what the instrument records is the cadence of the phrase superimposed on it.

One is an addition. At 20.4 the instrument designates a nucleus on position 3 that neither the lexicon nor the ear proposes. This is the criterion working as intended: it is free to name a syllable that no other reading predicts, and here it does.

None of these six results revises the annotation, in accordance with the rule of non-revision. What they establish is narrower and worth stating exactly: the categories counted in section 5 have measurable physical correlates in every verse measured, and the fundamental frequency alone locates those correlates one syllable away from the ear in half the sample. The accentual tradition since Werner has located its beats by ear, most probably in performance, where the most salient pitch movements belong to the melody rather than to the words. An instrument locates them on the raw contour, where pitch alone, as this sample shows, often falls one syllable away from the stress. Neither is a reliable guide to where a Swahili stress lies, which is why no instrumental reading in this study is permitted to decide on its own.

9.4.3. Conclusion on Utenzi Meter

Two corpora have been examined in the same meter, under the same protocol, by the same reader. The *Utenzi wa Rasi 'lGhuli* was composed at Bagamoyo and its language belongs to the southern dialectal cluster; the *Utendi wa Haudaji* belongs to the northern dialectal cluster. Three hundred and twenty lines were mapped, thirty removed for reasons of execution, and **290 lines** retained. The figures below are those of the two case studies, restated together, and each condition is closed with the verdict its figures support. No broader theoretical adjudication is attempted in this technical conclusion on the *utenzi* meter; what follows is strictly an account of what was measured and of the material facts that follow from it.

1. The syllabic model

Condition 1. The syllabic constant N = 8 holds in **282 of the 290 retained lines, that is 97.2 percent**. The eight exceptions are *guni* of the source texts, three in one corpus and five in the other, verified against the printed editions and retained in the pools, where they count against the matrix. No exception arises from the execution:

every line whose count departs from the source was removed before counting, and the removals are listed line by line in each case study.

Verdict. The condition is satisfied. Deviation is rare and every deviant line is *guni* in the endogenous sense, which is to say defective by the criterion the tradition itself states.

Condition 2. The global constant holds in **80 stanzas of 80**. It resolves differently in the two corpora: X^{s} in the *Rasi 'lGhuli*, where every anchored line closes on the syllable *-ri*, and X^{v} in the *Haudaji*, where X^{s} is contradicted by six anchored lines and the terminal vowel /a/ is what the poem holds constant. The local rhyme holds in **77 stanzas of 80** in its syllabic form and in 80 of 80 in its vocalic form.

Verdict. The condition is satisfied, and the hierarchy $X^{s} > X^{v}$ does the work it was defined to do: it records a real difference between two poems in one meter rather than accommodating both by relaxation.

Conditions 3 and 4 are not applicable to this meter, which has no internal caesura, nor to these corpora, which employ no permutation or stanzaic chaining.

The syllabic matrix therefore describes 97.2 percent of the lines exactly, accounts for the remaining 2.8 percent through a category the tradition itself defines and sanctions, and predicts the rhyme structure of both poems without exception. It should be stated that this is a verification and not a falsification: the matrices were formalized from the descriptions of the poets and applied to texts those poets did not analyze, and a line that contradicted them would have obliged them to change. None did, beyond the *guni*.

2. The accentual model

2.a. Is the bipartite axiom verified?

The lexicon of the line yields exactly two accentual groups in **161 lines of 290, that is 55.5 percent**. It yields three in **119 lines, that is 41.0 percent**, four in nine, and one in a single line. The formal reading realizes exactly two prominences in **175 lines, that is 60.3 percent**, and three in **106, that is 36.6 percent**.

Verdict. **The axiom is not verified.** An axiom admits no exception, and this corpus supplies 129 counter-examples on the lexical measure and 115 on the auditory one. The lines that contradict it are not a residue: two lines in five carry three accentual groups on the lexical measure, and more than one in three on the auditory one. A structure present in 41 percent of a corpus is not an irregularity to be set aside; it is a second outcome of the same frequency class as the first.

The transition between the two measures does not run toward the axiom. Of the 161 lines whose lexicon already yields two groups, 153 are realized with two. Of the 128 lines whose lexicon yields another number, **21 are brought to two in the reading, that is 16.4 percent**; 107 are not. The rise from 55.5 to 60.3 percent is produced by twenty-one lines across two poems, and the reading leaves 107 lines with a number of groups the axiom forbids.

The prose baseline. The reading departs from the lexical structure in **33 lines of 290, that is 11.4 percent**. In the control corpus of section 9.7 — eighty lines of ordinary Swahili prose read by the same speaker under the same instruction — the departure rate is **12.5 percent**, which is the higher of the two. A reader alters the accentual structure supplied by the lexicon at the same rate whether she is reading a nineteenth-century *utenzi* or a sentence about the price of fish. Whatever the transition from G-lex to G-aud measures, it is not an effect of meter.

2.b. Is the *Grenzschema* a cognitive prototype?

The claim is that $P_1 = 3$ and $P_2 = 7$ constitute an ideal configuration toward which the composer works. Four measurements bear on it.

Frequency. The lexical sequence 4·4 occurs in **49 lines of 290, that is 16.9 percent**, and is realized as $P_1 = 3$, $P_2 = 7$ in **56 lines, that is 19.3 percent**. The complete model — the correct number of groups *and* the correct placement — is therefore satisfied by fewer than one line in five.

Rank. Among the 175 lines that satisfy the bipartite axiom, the two prominences fall on positions 2 and 7 in **60 lines**, on 3 and 7 in **56**, on 4 and 7 in **35**, on 1 and 7 in **16**, and on five further configurations in one to three lines each. The configuration designated as the prototype **is not the most frequent even within the set of lines most favorable to it**: the sequence 3·5, with the first prominence on position 2, occurs more often. Of the lines that vindicate the model on the number, 32.0 percent vindicate it on the placement.

The broadening of the criterion, and its failure. Minerba does not restrict the trochaic rhythm to the *Grenzschema* alone. He admits three of the five possible dispositions of the two stresses as

trochaic, on the ground that they place the accents on odd-numbered syllables, and excludes the other two:

> The two forms σ´ σ | x x x x σ´ σ and x x x x σ´ σ | σ´ σ can also be considered to convey a trochaic rhythm: in this case too, stresses fall on odd positions, therefore the rhythmic impression of trochee is created, even if in a weaker manner. The forms x σ´ σ | x x x σ´ σ and x x x σ´ σ | x σ´ σ are, instead, allowed variation which do not form the trochaic rhythmic impression: being closer to the archaic patterns than to the classical ones, their employment is nevertheless permitted to avoid having an over-strict reproduction of the *Grenzschema*. (Minerba, 2025, p. 220)

The criterion is therefore the parity of the first stressed position: odd for trochaic, even for the permitted variation. Tested on that criterion, the model fares no better and the outcome is reversed against it. Of the 161 lines whose lexicon yields two accentual groups, **70 place the first stress on an odd syllable, that is 43.5 percent**, and **91 place it on an even syllable, that is 56.5 percent**. Of the 175 lines realized with two prominences, **77 are trochaic by this criterion, that is 44.0 percent**, and **98 are not, that is 56.0 percent**. The two corpora agree: 41.9 percent of trochaic lines in one, 46.1 percent in the other.

The majority of the lines that satisfy the bipartite axiom therefore fall outside the rhythm the axiom is invoked to explain, and they do so by the definition of the scholar who proposes it. The most frequent single disposition across both corpora places the first stress on position 2, which is one of the two forms he classes as non-trochaic.[86]

---

[86] This outcome was available to Minerba on his own sample. Counting the hemistiches of Muyaka's *Kimya*, he arrives at a near-equal division and interprets it as an intended effect: "In Kimya, trochaic and non trochaic hemistichs follow each other with almost the same frequency (13 hemistichs are trochaic, 11 are not); this gives movement and rhythmic complexity to the composition." (Minerba, 2025, p. 220)

The reasoning should be examined, because it determines what could count against the model. A schema realized in thirteen segments and contradicted in eleven is not thereby confirmed; it is a distribution close to even. Presenting the contradicting half as rhythmic complexity converts a negative result into a positive one without introducing any new evidence, and leaves no distribution capable of falsifying the claim. On the corpora measured here the division is not even but adverse, 98 against 77, and the same reasoning would have to describe the majority of the lines as the complexity of a rhythm that the minority establishes.

Stability across works. The sequence occurs in 11.4 percent of the *Haudaji* and in 22.7 percent of the *Rasi 'lGhuli*, a ratio of two to one between two poems in the same meter. A generative matrix does not vary by a factor of two from one work to another in the same form; that is what makes it a matrix. What varies here is the lexicon each composer drew on.

Presence outside poetry. The eight lines of the prose control corpus whose lexicon yields the sequence 4·4 are realized with $P_1 = 3$ and $P_2 = 7$ in all eight cases, and no line whose lexicon yields another sequence is so realized. The configuration arises automatically, in prose written for no metrical purpose, whenever the lexicon supplies two words of four syllables.

Verdict. **The *Grenzschema* is not a cognitive prototype.** A prototype is a configuration toward which production tends. This one occurs in fewer than a fifth of the lines, ranks second within the set most favorable to it, varies by a factor of two between two works in the same meter, and appears in prose at the rate the lexicon dictates and no other. What the measurements describe is not an attractor but an arithmetical consequence: the sequence occurs when, and only when, the composer's words happen to fall into two groups of four syllables.

2.c. Is there a rhythm?

A rhythm requires a unit that repeats. The question is therefore asked at the two scales at which repetition could occur, and answered separately at each.

Within the line. A line exhibits an internal periodicity only if its accentual nuclei are of equal volume. This holds in **59 lines of 290, that is 20.3 percent** — forty-nine in the sequence 4·4, nine in 2·2·2·2, one in 3·3·3. In the remaining **231 lines, that is 79.7 percent**, the nuclei are of unequal size and no unit repeats. Among the lines that yield two groups, the attested sequences are 4·4, 3·5, 5·3, 2·6, 6·2, 6·3, 3·4, 4·3 and 5·4; the first stressed syllable falls on position 1, 2, 3, 4 or 5, and the interval separating the two stresses varies from two to six syllables.

Across lines. Two successive lines carry the same sequence of nucleus volumes in **42 pairs of 288, that is 14.6 percent**, against **37.5 pairs expected under independence, that is 13.0 percent** — an excess of fewer than five pairs in two hundred and eighty-eight. Within the stanza, where a compositional effect would be most expected, 34 pairs of 211 are identical, that is 16.1 percent. Of the seventy-five stanzas whose four lines are all retained, **one** carries the same sequence on all four.

Verdict. **There is no rhythm at either scale.** Four lines in five contain no repeating unit, so that no foot can be identified within them. The succession of configurations from one line to the next is statistically indistinguishable from independent draws, so that no periodicity is established between them. This result requires no recording and no listener: it is computed from the printed lines by the penultimate rule, and it holds even if the bipartite axiom were granted in full. A line may carry exactly two prominences and still contain nothing that repeats.

Is the tetrasyllabic template used?

The *utenzi* cannot falsify Condition 8 arithmetically, since eight is a multiple of four and the template is available in every regular line. What the corpora establish is whether it is taken up. A tetrasyllabic nucleus of any kind occurs in **92 lines of 290, that is 31.7 percent**, and only **49 lines** are partitioned as two tetrasyllabic units. In 43 lines the four-syllable group is a single long word flanked by shorter ones, and in **198 lines, that is 68.3 percent**, no tetrasyllabic nucleus occurs at all.

Verdict. The template is available everywhere and taken up in fewer than one line in six. The condition is not decisive for this meter and is tested where it can be, in the *wimbo* corpora.

The instrumental sample

Sixteen lines were measured, eight in each corpus, on samples whose composition was fixed before measurement, and which include the ground most favorable to the accentual model. The instrumental reading coincides exactly with the auditory annotation in **five lines of sixteen**.

Three observations follow from the eleven departures. Prominence on position 7 is confirmed in every line where the ear placed it, save one. The departures are predominantly displacements of a single syllable, produced by the delay between the onset of a syllable and the culmination of the melodic gesture it initiates, measured between 53 and 221 milliseconds across the sample. And in one line the two procedures return incompatible results: read by ear, *kutunga kwa ushairi* carries its prominences on positions 2 and 7; read by melodic salience, on positions 1 and 8.

Verdict. Both readings are readings of the same performance, obtained by procedures the accentual tradition would accept, and they do not agree. There is no single fact of the matter about where the beats of a line lie, even within one reading by one speaker measured one way. The syllable count admits no such indeterminacy.

The origin of the disagreement is worth naming, since it bears on the tradition under test rather than on the sample. Where a melodic rise is taken as the mark of a beat, what is being recorded is the shape of the contour and not the structure of the line: the peak arrives after the syllable that produced it, the opening of a line resets the register, and its close carries a boundary tone. A model built on that mark therefore imports the physiology of speech into its inventory of metrical facts, and counts as constitutive of the verse what is in truth a property of any Swahili utterance delivered aloud.

What the two models account for

The syllabic matrix predicts the number of syllables in 97.2 percent of the lines and the rhyme structure in 100 percent of the stanzas, from the printed text, without a recording, by a procedure any reader may repeat.

The accentual matrix predicts the number of accentual groups in 55.5 percent of the lines and, where the number is correct, the placement in 32.0 percent of them: the complete model holds in **16.9 percent of the corpus** on the lexical measure and 19.3 percent on the auditory one. It provides no recurrent unit in 79.7 percent of the lines, no sequencing from line to line beyond what independence would produce, and no configuration that the prose control does not produce at the same rate under the same conditions. Its single invariant, the prominence on position 7, follows from the penultimate rule applied to the final word of a line that admits no enjambment, and is present identically in prose.

Of the four conditions that test the accentual model, Condition 5 is contradicted by 129 lines, Condition 6 by the rank, the instability and the prose baseline of the prototype, and Condition 7 by the absence of any repeating unit in four lines out of five and by the independence of consecutive lines. Condition 8 is not decisive for this meter.

On the evidence of these two corpora, the accentual model of the *utenzi* line is not confirmed on any condition under which it could have been. What it describes — prominence, and its frequent occurrence in pairs — is present in the data, and is present in the same proportion in prose composed for this study with no metrical intention whatever. The accentual configurations of the *utenzi* line are therefore a property of the Swahili lexicon within an eight-syllable frame, and not a property of the meter.

The same conditions are applied to the *wimbo* and to the *shairi* in the case studies that follow, and the three partial results are drawn together in the conclusion of this book.

9.5. *Wimbo* Texts

9.5.1. Case Study III: Test of a Rupture Matrix — Wimbo wa Miti — Woman Author Unknown (c. 1800–1850)

To conclusively test the explanatory power of the endogenous syllabic grid against the exogenous hypotheses of the accentual school, our empirical analysis must deliberately venture into the metrical margins of the classical corpus. The text subjected to this clinical verification is the *Wimbo wa Miti* (The Song of the Trees) (Mbele, 1996). Collected in June 1989 by the researcher Joseph L. Mbele from Zaharia Maimun—an eminent Swahili poetess and scholar residing in Lamu—this text provides an invaluable window into the complex mechanics of female-authored Swahili verse.

1. Presentation, Dating, and Philological Context

The *Wimbo wa Miti* is an allegorical poem functioning strictly as a *zifungo* (an enigma poem or intellectual riddle). Zaharia Maimun insisted that this composition belongs to the *wimbo* meter, explicitly criticizing the common academic and popular tendency to inaccurately label all Swahili poetry as *shairi* or *utendi*.

Regarding its historical origin, Maimun stated that the poem is "very old, much older than the time of her own grandmother" (Mbele, 1996). Given that Maimun was approximately sixty years old at the time of the 1989 interview (placing her birth circa 1929), her grandmother's era would securely fall in the late 19th century. The designation "much older" anchors the composition of this text in the first half of the 19th century (circa 1800–1850). The poem therefore belongs entirely to the classical golden age of the Lamu archipelago.

According to Mbele (1996), the genesis of the poem is rooted in social mediation. It was composed by an anonymous woman to cryptically address and rectify severe misunderstandings, backbiting, and malicious gossip that had fractured the community of women in Lamu town. By utilizing a sophisticated allegorical mode where trees and plants embody specific human temperaments—the furious mango tree representing anger, the pepper plant representing blind aggression, and the pumpkin plant serving as the hidden agitator—the poetess formulated a devastating social critique without ever directly naming the belligerents. The song was traditionally performed at

weddings, functioning as a didactic intellectual challenge for the audience and fellow poets to decode. It stands as a testament to Swahili female heroism, advocating for social cohesion and harmony in stark contrast to the destructive, warlike values often celebrated in the male-dominated epic poetry.

2. The Rupture Matrix: The Geometric Architecture of the Wimbo

Minerba postulates that the classical Swahili meters — *utenzi*, *shairi* and *wimbo* — are the product of a trochaic innovation inspired by Arabic vernacular meters and based on tetrasyllabic feet, that is on multiples of four (Minerba, 2025, pp. 217–219). He argues that the internal caesuras of the *wimbo*, yielding forms such as 8 + 8, 4 + 8 or 8 + 4, were engineered to isolate these four-syllable blocks and thereby force a constant underlying trochaic rhythm (Minerba, 2025, p. 221). The formalization of this claim, its two degrees of strictness and its conditions of failure are set out in section 2.5.2; the conditions themselves are stated in section 9.2.4, and it is Condition 8 that this case study puts to the test.

To insulate the theory from the reality of the manuscripts, Minerba establishes a statistical framework grouping his measurements into three categories according to "their adherence to the theorized norm: fulfilment, exception, and error" (Minerba, 2025, p. 101). Deviations from the expected model — a line of eleven syllables in a sequence of twelve — are dismissed as errors (Minerba, 2025, p. 102), while structurally inconvenient nine-syllable hemistiches are explained as exceptions relying on prosodic contractions (Minerba, 2025, p. 103). By marginalizing the asymmetrical, non-tetrasyllabic matrices he himself lists — 6 + 6, 2 + 8, 6 + 4, 8 + 5 (Minerba, 2025, p. 160) — and by filtering structural variation as anomaly, the framework resists empirical verification a priori. A model that validates only the data fitting a multiple of four, and discards the remainder as defective or exceptional, has ceased to be an observation and become a procedure immunized against falsification.

The *Wimbo wa Miti* contradicts the theory on the ground where Condition 8 is decisive, and it does so before any question of prominence arises. The poem is structured in tercets, $n_{max} = 3$, and deploys a vertical asymmetry: the caesura does not remain static but shifts within the stanza itself.

This architecture answers a strict algebraic equation anchored in Layer 1, the syllabic law. We defined this metrical current (*mkondo*) as the Rupture Matrix in section 8.4.

Formal Parameters:

The poem is a tercet ($n_{max} = 3$).

The line is divided into two hemistiches ($H = 2$).

The variable of the asymmetrical first hemistich is established at $N = 2$ syllables in this poem.

The poem systematically alternates lines of 2 + 8 syllables with a final, massive line of 8 + 8 syllables.

The General Matrix:

L_1: [N, a] + [8, b] //
L_2: [N, a] + [8, b] //
L_3: [8, a] + [8, b] //

The Algebraic Equation for the Rupture *Wimbo*:

$$\boldsymbol{Wimbo_{rupture} = 2(N_a + 8_b//) \cdot (8_a + 8_b//)}$$

With N = 2 in this poem.

3. Experimental Verification

The verification follows the three-step protocol set out in section 9.2.3, and it is stated here so that no step is credited with more than it performs.

The accentual structure of every line is first derived from the printed text by the penultimate rule. This derivation requires no recording, no listener and no software, and any reader who knows the rule recomputes it from the text as printed. It is the foundation of what follows.

The formal reading is then screened line by line by ear, and the two results are reported side by side for the whole poem. Acoustic extraction is applied to a sample of lines only, and no statistic in this case study rests upon it.

The reading is a reading and not a performance, which sets aside Layer 3. What the reader receives is the page itself: the graphic manifestation of the syllabic count and the rhyme boundaries of Layers 1 and 2, carrying no accentual notation of any kind. The accentual structure is supplied by the language and not by the page — Swahili assigns one stress to each phonological word, on its penultimate syllable, so that a given sequence of words determines one accentual structure and only one. That structure is not a specification of Layer 1, which is silent on accent; it is a property of Swahili which the accentual school has taken for a metrical law.

Layer 3 is set aside because it is not deterministic. There the melody (*mahadhi*) and musical time distort its realization: a line whose words determine one accentual structure can yield an unlimited number of conflicting realizations depending on the singer's choices, which stretch vowels, introduce non-lexical rests, and displace or ignore lexical stresses to fit a tune. The formal reading realizes only part of the lexical structure, and the two are therefore recorded as two distinct datasets throughout.

4. Wimbo wa Miti: Syllabic and Accentual Mapping, Stanzas 1–8

4.a. Reading the entries

Each line occupies a block of three short lines, in the notation of sections QRG.1 and 2.5. The first gives the reference and the line as printed, syllabified, with the two hemistiches separated by the caesura marker + and their syllable counts *(mizani)* in brackets.

The second is the lexical reading, derived from the printed text by the penultimate rule and requiring no recording: for each hemistich, the sequence of nucleus volumes $|A_1|\cdot|A_2|\cdot\ldots$, the positional index P*i* of the stressed syllable of each nucleus, and G-lex. Positions reset at the caesura. The third is the auditory reading, established by ear on the formal reading, followed by Δ, which records the transition for each hemistich.

What is observed is whether a tetrasyllabic rhythm survives this asymmetrical grid, and — where the grid permits the template at all — whether the first stressed position falls on an odd syllable, which is the extended criterion Minerba (2025) states at page 220 and which section 2.5.2 formalizes.

A question mark marks a weak realization. On a nucleus volume it indicates a nucleus that is counted although its prominence is markedly lower than that of the others in the same hemistich. On a deletion in the Δ column it indicates a lexical nucleus judged

unrealized, the judgment resting on a prominence too faint to count rather than on its absence. In both cases the mark records a property of the reading, not a hesitation of the listener, and the decision stands as written. No such mark can appear in the lexical reading, which is derived by rule.

4.b. Mapping

Stanza 1 — rhyme a = *na*, rhyme b = *mu*
1.1 Ya-na + ni-we-ne ki-sa a-dhi-mu (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·2?·3 — P1=1 | P1=2, P2=4?, P3=7 — G-aud 1 + 3 — Δ — | —
1.2 te-na + ta-fu-si-ri-ni wa-li-mu (2+8)
lexical 2 + 5·3 — P1=1 | P1=4, P2=7 — G-lex 1 + 2
auditory 2 + 5·3 — P1=1 | P1=4, P2=7 — G-aud 1 + 2 — Δ — | —
1.3 Mi-ti ya-li-ki-pi-ja-na + m-te-nde-ti na m-di-mu (8+8)
lexical 2·6 + 4·4 — P1=1, P2=7 | P1=3, P2=7 — G-lex 2 + 2
auditory 2·6 + 4·4 — P1=1, P2=7 | P1=3, P2=7 — G-aud 2 + 2 — Δ — | —
Stanza 2 — rhyme a = *ma*, rhyme b = *ra*
2.1 Hi-ma + u-ka-ya mbi-yo m-pwe-ra (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·2?·3 — P1=1 | P1=2, P2=4?, P3=7 — G-aud 1 + 3 — Δ — | —
2.2 Nyu-ma + mu-e-mbe u-na ha-si-ra (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·2?·3 — P1=1 | P1=2, P2=4?, P3=7 — G-aud 1 + 3 — Δ — | —
2.3 M-za-bi-bu u-ki-se-ma + ku-nya-ma-a ni u-ju-ra (8+8)
lexical 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G-lex 2 + 2
auditory 4·4 + 4?·4? — P1=3, P2=7 | P1=3?, P2=7? — G-aud 2 + 2 — Δ — | —
Stanza 3 — rhyme a = *ra*, rhyme b = *si*
3.1 Ma-ra + u-ka-dhi-hi-ri m-te-si (2+8)
lexical 2 + 5·3 — P1=1 | P1=4, P2=7 — G-lex 1 + 2
auditory 2 + 5·3 — P1=1 | P1=4, P2=7 — G-aud 1 + 2 — Δ — | —
3.2 su-ra + zi-na u-shi-u na ka-si (2+8)
lexical 2 + 2·3·3 — P1=1 | P1=1, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 2·3·3 — P1=1 | P1=1, P2=4, P3=7 — G-aud 1 + 3 — Δ — | —
3.3 U-ki-wa-mbi-ya m-pwe-ra + kwe-tu ha-ku-na ra-khi-si (8+8)
lexical 5·3 + 2·3·3 — P1=4, P2=7 | P1=1, P2=4, P3=7 — G-lex 2 + 3
excluded — false start on the first nucleus of H1

Stanza 4 — rhyme a = *po*, rhyme b = *li*
4.1 Pa-po + u-ka-ya m-pi-li-pi-li (2+8)
lexical 2 + 3·5 — P1=1 | P1=2, P2=7 — G-lex 1 + 2
auditory 2 + 3·3·2 — P1=1 | P1=2, P2=5, P3=7 — G-aud 1 + 3 — Δ — | +P2=5
4.2 Ha-po + u-ki-ne-na kwa u-ka-li (2+8)
lexical 2 + 4·4 — P1=1 | P1=3, P2=7 — G-lex 1 + 2
auditory 2 + 4·4 — P1=1 | P1=3, P2=7 — G-aud 1 + 2 — Δ — | —
4.3 U-ki-a-pa kwa zi-ya-po + mi-ti pi-a si-i-ja-li (8+8)
lexical 4·4 + 2·2·4 — P1=3, P2=7 | P1=1, P2=3, P3=7 — G-lex 2 + 3
auditory 4·4 + 2·2·4 — P1=3, P2=7 | P1=1, P2=3, P3=7
— G-aud 2 + 3 — Δ — | —
Stanza 5 — rhyme a = *nde*, rhyme b = *o*
5.1 Pu-nde + wa-si-ye-sa ka-ni za-o (2+8)
lexical 2 + 4·2·2 — P1=1 | P1=3, P2=5, P3=7 — G-lex 1 + 3
auditory 2 + 4·2·2 — P1=1 | P1=3, P2=5, P3=7 — G-aud 1 + 3 — Δ — | —
5.2 n-de + wa-we-ne m-ti u-ya-o (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-aud 1 + 3 — Δ — | —
Note. The prefix *n* in *nde* (standard *nje*) functions as a syllabic nasal, and it is what allows the hemistich to meet the two-syllable requirement of the matrix exactly.
5.3 Na-o ni m-be-re-me-nde + u-na pa-nga na ma-nga-o (8+8)
lexical 2·6 + 2·2·4 — P1=1, P2=7 | P1=1, P2=3, P3=7 — G-lex 2 + 3
auditory 2·4·2 + 2?·2·4 — P1=1, P2=5, P3=7 | P1=1?, P2=3, P3=7
— G-aud 3 + 3 — Δ +P2=5 | —
Stanza 6 — rhyme a = *ta*, rhyme b = *wa*
6.1 Zi-ta + ha-tu-chi twa-po u-wa-wa (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-aud 1 + 3 — Δ — | —
6.2 U-ta + na ze-mbe tu-me-tu-ku-wa (2+8)
lexical 2 + 3·5 — P1=1 | P1=2, P2=7 — G-lex 1 + 2
auditory 2 + 3·5 — P1=1 | P1=2, P2=7 — G-aud 1 + 2 — Δ — | —

6.3 M-bu-yu u-li-po-pi-ta + u-ka-si-ma-ma kwa mu-wa (8+8)
lexical 3·5 + 5·3 — P1=2, P2=7 | P1=4, P2=7 — G-lex 2 + 2
auditory 3?·5 + 5·3 — P1=2?, P2=7 | P1=4, P2=7 — G-aud 2 + 2 — Δ — | —
Note. Mbele translates the final word *muwa* as "mediate". In standard Swahili this is *kuamua* (ku-a-mu-a), four syllables. The poetess contracts it to *muwa* (mu-wa), two syllables, which combined with *kwa* (one) and *ukasimama* (five) yields exactly eight. The morphology is adjusted to meet the syllabic count, and not the reverse.
Stanza 7 — rhyme a = *la*, rhyme b = *we*
7.1 He-la + wa-ta-ni ye-nu ma-yo-we (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·5 — P1=1 | P1=2, P2=7 — G-aud 1 + 2 — Δ — | –P2=4?
7.2 te-la + bu-re a-si-wa-zu-zu-we (2+8)
lexical 2 + 2·6 — P1=1 | P1=1, P2=7 — G-lex 1 + 2
auditory 2 + 2·6 — P1=1 | P1=1, P2=7 — G-aud 1 + 2 — Δ — | —
7.3 Ku-na m-ti u-na hi-la + hu-m-shi-nda ki-tu-ngu-we (8+8)
lexical 2·2·2·2 + 4·4 — P1=1, P2=3, P3=5, P4=7 | P1=3, P2=7 — G-lex 4 + 2
auditory 2·2·4 + 4·4 — P1=1, P2=3, P3=7 | P1=3, P2=7 — G-aud 3 + 2 — Δ –P3=5? | —
Stanza 8 — rhyme a = *na*, rhyme b = *wa*
8.1 Shi-na + la ma-mbo ha-ya na-yu-wa (2+8)
lexical 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-lex 1 + 3
auditory 2 + 3·2·3 — P1=1 | P1=2, P2=4, P3=7 — G-aud 1 + 3 — Δ — | —
8.2 Sa-na + ni-me-zi-ye ku-ta-mbu-wa (2+8)
lexical 2 + 4·4 — P1=1 | P1=3, P2=7 — G-lex 1 + 2
auditory 2 + 4·4 — P1=1 | P1=3, P2=7 — G-aud 1 + 2 — Δ — | —
Note. Mbele's transcription gives the first word as *Sawa*. The generative matrix requires the local rhyme *a* to be *na*, matching *Shi-na* at 8.1 and *fi-ti-na* at 8.3. The syllabic grid identifies *Sawa* as a transcription error and confirms that the word must be *Sa-na*.
8.3 Mu-yu-ngu ndi-o fi-ti-na + wa-lo-a-li-sha ma-u-wa (8+8)
lexical 3·2·3 + 5·3 — P1=2, P2=4, P3=7 | P1=4, P2=7 — G-lex 3 + 2
auditory 3·2?·3 + 5·3 — P1=2, P2=4?, P3=7 | P1=4, P2=7
— G-aud 3 + 2 — Δ — | —

4.c. Excluded lines

One line is removed from the statistical pool. Line 3.3 was begun twice by the reader, and a false start on the first nucleus of its opening hemistich leaves the realization of that line unverifiable; it is excluded

entire, as in the *utenzi* corpora, since a hemistich delivered after a restart cannot be set beside one delivered in a single pass.

3.3 U-ki-wa-mbi-ya m-pwe-ra + kwe-tu ha-ku-na ra-khi-si — false start on the first nucleus of H1

4.d. Summary

The pool comprises **23 lines of the 24 mapped**, that is **46 hemistiches**: sixteen of two syllables and thirty of eight. The syllabic constants hold without exception: every line of the first two positions of each stanza divides as 2 + 8, and every third line as 8 + 8, with no deviation in the twenty-three retained lines.

The two-syllable hemistiches. All sixteen are realized with a single accentual nucleus of two syllables, without exception. The measure leaves the penultimate rule no other outcome, and the reading produces none.

The eight-syllable hemistiches. Distribution of G-lex, derived from the printed lines: G = 2 in 17 hemistiches (56.7 percent), G = 3 in 12 (40.0), G = 4 in 1 (3.3). Distribution of G-aud: G = 2 in 16 (53.3 percent), G = 3 in 14 (46.7).

The reading departs from the lexical structure in **4 of the 30 eight-syllable hemistiches, that is 13.3 percent**. The departures are listed individually: at 4.1 the sequence 3·5 is realized as 3·3·2; at 5.3 the sequence 2·6 is realized as 2·4·2; at 7.1 the sequence 3·2·3 is realized as 3·5; at 7.3 the sequence 2·2·2·2 is realized as 2·2·4.

The sequence 4·4 occurs in **7 hemistiches (23.3 percent)** on the lexical measure and in **7 (23.3 percent)** on the auditory one. The most frequent lexical sequence of this corpus is not 4·4 but 3·2·3, which occurs in eight hemistiches, a sequence of three nuclei of unequal volume.

Nine of the thirty eight-syllable hemistiches, that is 30.0 percent, carry at least one nucleus marked as weakly realized.

5. Statistical Analysis and Falsification

5.a. What is tested here, and what is not

This corpus is not equivalent to an *utenzi* corpus, and the conditions of section 9.2.4 cannot be applied to it in the same terms. The difference must be stated before any figure is given, since it governs what the figures can establish.

The *utenzi* line is a single undivided segment of eight syllables, and every line of a forty-stanza corpus offers the same container to the same test. The *Wimbo wa Miti* offers three different containers. Its first two lines divide as 2 + 8 and its third as 8 + 8, so that a stanza yields two hemistiches of two syllables and four of eight. The unit of analysis is therefore the hemistich and not the line, and the two lengths are not tested for the same thing.

The two-syllable hemistiches are not tested for accentual structure at all. A segment of two syllables admits exactly one accentual nucleus, the penultimate rule having no second position to occupy. Whatever the reader does, G can only be 1. Counting these sixteen hemistiches among the outcomes of Condition 5 would inflate the corpus with a result that no reading could contradict. They are reported in section 4 and set aside here, and their behavior is examined instead under Condition 8, where a segment too short to host a tetrasyllabic foot is precisely what is at issue.

The eight-syllable hemistiches are tested exactly as *utenzi* lines are. They are containers of eight syllables filled by the Swahili lexicon under a metrical constraint, which is what Conditions 5, 6 and 7 examine, and it is on these thirty hemistiches that the accentual model is put to the test. The prose baseline of section 9.7 applies to them without adjustment, since it measures what the language does under a constraint of eight syllables, whatever meter imposes that constraint.

Conditions 1 and 2 are tested on the line, since the syllabic constants and the rhyme boundaries are properties of the whole line and of the stanza.

Condition 3 becomes applicable for the first time in this chapter, the *utenzi* having no internal caesura.

Condition 4 is not applicable, this corpus employing no permutation or stanzaic chaining.

5.b. Conditions 1 and 2: the syllabic matrix and the rhyme

Condition 1, the syllabic constants. The Rupture Matrix requires 2 + 8 in the first two lines of every stanza and 8 + 8 in the third. Across the twenty-three retained lines the constants hold **without a single exception**. No hemistich of the corpus deviates by one syllable from the value its position prescribes, and no *guni* of the source text occurs.

Two lines of the corpus show how that result is obtained, and they are worth reporting because they show the constraint operating on the

composer rather than describing her after the fact. At 5.2 the two-syllable hemistich is *n-de*, where the prefix *n* functions as a syllabic nasal and supplies the second syllable the matrix requires. At 6.3 the poetess renders what standard Swahili would give as *kuamua*, four syllables, by the contracted *muwa*, two syllables, which combined with *kwa* and *ukasimama* yields exactly eight. The morphology is adjusted to the count, and not the count to the morphology.

Condition 2, the rhyme variables. The matrix leaves both rhyme variables unresolved, and each is tested in the order its hierarchy prescribes. The local rhyme *a* binds the first hemistich of each line within the stanza; the rhyme *b* binds the second.

a^{s}, the syllabic form, is tested first and holds in **all eight stanzas** for both rhymes. Every *a* rhyme closes on the same full syllable throughout its stanza — *na*, *ma*, *ra*, *po*, *nde*, *ta*, *la*, *na* — and every *b* rhyme likewise — *mu*, *ra*, *si*, *li*, *o*, *wa*, *we*, *wa*. No stanza requires the vocalic form, and a^{v} is not invoked.

The result is more demanding than it appears. This matrix asks the poetess to rhyme a two-syllable hemistich with an eight-syllable one in the same stanza, on the same full syllable, and to do so eight times over. She does it without a single license.

Conditions 1 and 2 are therefore satisfied without exception. It should be stated plainly what this establishes and what it does not: these conditions verify the exhaustiveness of a description derived from Shariff (see Section 8.4.2.) and applied to a text he did not analyze; they do not falsify a rival hypothesis.

5.c. Condition 3: the status of the caesura

Condition 3 asks whether the natural syntax of the text respects the metrical boundary, or whether the composer forces a single morphological word across it. The criterion is decisive for the matrix: a caesura that words cross freely is a typographical convention and not a compositional constraint.

In the twenty-three retained lines, **no word crosses a caesura**. Every hemistich boundary of this corpus coincides with a word boundary, without exception, including the sixteen boundaries that follow a two-syllable hemistich, where the constraint is at its tightest.

The result acquires its force from the shape of the matrix. A caesura placed after two syllables leaves the composer almost no room: the opening hemistich must be a complete word or word group of exactly two syllables, carrying the stanza's *a* rhyme, and the following hemistich must begin a new word. She meets it sixteen times. The

boundary is not a convenience of the transcription; it is a constraint the composer accepted and satisfied.

5.d. Condition 5: the bipartite accentual axiom

The axiom requires that an eight-syllable segment carry exactly two accentual groups. The thirty eight-syllable hemistiches of this corpus are tested against it, on the lexicon and on the reading.

The lexicon yields exactly two accentual groups in **17 hemistiches of 30, that is 56.7 percent**, three in 12 (40.0 percent), and four in one. The reading realizes two in **16 (53.3 percent)** and three in 14 (46.7 percent).

Verdict. The axiom is not verified. It is contradicted by 13 hemistiches on the lexical measure and by 14 on the auditory one, and the configuration it forbids — three accentual groups — occurs in two hemistiches of every five.

The transition. The reading departs from the lexical structure in **4 hemistiches of 30, that is 13.3 percent**. In the prose control corpus of section 9.7, read by the same speaker under the same instruction, the rate is 12.5 percent. The two figures are of the same order, and the difference between them amounts to a quarter of one hemistich in thirty. Whatever the departure measures, it is not far from what ordinary prose produces under the same syllabic constraint.

One qualification belongs here rather than in a footnote. This corpus is a *wimbo*, a sung form, and nine of its thirty eight-syllable hemistiches carry a nucleus marked as weakly realized, which is the highest proportion recorded in this chapter. The reading may lean further toward the melodic execution of Layer 3 than an *utenzi* reading does, and thirty hemistiches are too few to settle the question. The figures are reported as they stand.

5.e. Condition 6: the *Grenzschema*

The ideal configuration requires $|A_1| = 4$, $|A_2| = 4$ with stresses on positions 3 and 7. It occurs in **7 hemistiches of 30 on the lexical measure, that is 23.3 percent**, and in **7 on the auditory one, that is 23.3 percent**. The reading neither adds an instance nor removes one.

Rank. The most frequent lexical sequence of this corpus is not 4·4 but **3·2·3, with eight occurrences against seven** — three nuclei of unequal volume, which Minerba classes outside the trochaic

rhythm. Among the sixteen hemistiches that satisfy the bipartite axiom, the *Grenzschema* accounts for seven, that is 43.8 percent.

The extended parity criterion. Minerba admits as trochaic any disposition placing the first stress on an odd syllable (Minerba, 2025, p. 220). Of the sixteen hemistiches realized with two accentual groups, **9 satisfy that criterion, that is 56.2 percent**, and 7 do not. The first stress falls on position 3 in seven hemistiches, on position 4 in four, on position 2 in three and on position 1 in two.

Verdict. On the extended criterion the model holds for a majority of the two-group hemistiches, and this should be said: it is the first corpus of this chapter where it does. On the strict criterion it governs fewer than one hemistich in four, and it is outranked in the lexicon by a sequence it classes outside the rhythm.

5.f. Condition 7: periodicity

A foot is a unit that repeats. A segment exhibits an internal periodicity only if its accentual nuclei are of equal volume. This holds in **8 hemistiches of 30, that is 26.7 percent** — the seven sequences 4·4 and the single 2·2·2·2, and nothing else. In the remaining 22 the nuclei are of unequal size and no unit repeats.

The sequences attested on the lexical measure are 3·2·3, 4·4, 5·3, 2·6, 3·5, 2·2·4, 2·3·3, 4·2·2 and 2·2·2·2. Among the seventeen hemistiches with two groups on the lexical measure, the interval separating the two stresses ranges from four to six syllables and the first stress falls on positions 1 to 4.

Across hemistiches. Two consecutive eight-syllable hemistiches carry the same sequence of nucleus volumes in **4 pairs of 29, that is 13.8 percent**, against 4.5 pairs expected under independence, that is 15.6 percent. The observed frequency is below the expected one.

Verdict. There is no rhythm at either scale. Three hemistiches in four contain no repeating unit, and the succession of configurations from one hemistich to the next is not distinguishable from independent draws — indeed it falls slightly below what independence predicts.

5.g. Condition 8: the tetrasyllabic foot

This is where the *wimbo* differs from the *utenzi*, and the difference is arithmetical rather than statistical.

Of the **46 hemistiches** of the retained corpus, **16 measure two syllables**, that is 34.8 percent. A segment of two syllables cannot host

a four-syllable foot, and no partition rescues it. For more than a third of the metrical units of this poem, Minerba's derivation does not predict badly: it produces no prediction at all, while continuing to be asserted of the meter that contains them (Minerba, 2025, p. 217).

Where the template is arithmetically available, in the thirty eight-syllable hemistiches, it is used in **10 of them, that is 33.3 percent**, and only 7 are partitioned entirely into tetrasyllabic units. In twenty hemistiches, that is 66.7 percent, no tetrasyllabic nucleus occurs at all.

Verdict. The template is unavailable in a third of the corpus and unused in two thirds of the remainder. And the sixteen segments that make it unavailable are not defects of transmission: they are the constitutive feature of this *mkondo*, repeated on every first and second line of every stanza, each locked by its own rhyme. The composer builds the boundary that forbids the foot.

9.5.2. Case Study IV: *Wimbo* — Test of an Asymmetrical Bicolumnar Matrix (4 + 8) — The Accentual School's Home Ground

To continue the empirical verification of the classical Swahili metrical grid, we turn to the asymmetrical configurations of the *wimbo*. The textual data is drawn from Ibrahim Noor Shariff's *Tungo Zetu* (1988, pp. 45–49), where Shariff catalogues the structural diversity of the *wimbo* through a sequence of examples, labeled 'a' to 'k', documenting the various *mikondo* engineered by Swahili composers.

1. The Epistemological Value and Limits of the Fragment

Before deploying our formal models, we must establish a key methodological caveat regarding the nature of the data itself. Both Emiliano Minerba's accentual theories and our own data-driven syllabic formalizations rely heavily on these specific pages from Shariff's work. It is vital to recognize that the examples provided by Shariff—including example "b," which we will analyze here—are frequently fragments. They are isolated stanzas or partial sequences extracted from larger works, utilized by Shariff for purely pedagogical purposes to demonstrate the mechanical variations of the metrical grid.

From a strictly scientific and statistical perspective, relying on fragments is highly instructive but not fully representative of the macroscopic reality of complete, multi-stanzaic classical poems operating in their natural manuscript habitat. We are currently in a state of awaiting more comprehensive, unfragmented archival data to expand the corpus for large-scale analyses. However, for the precise purpose of testing the generative equation at the micro-stanzaic level—specifically verifying whether a line obeys an accentual beat or an absolute syllabic boundary—these fragments are perfectly valid and remain absolute primary evidence of the Swahili poetic intellect.

2. The Asymmetrical Matrix (4+8)

We focus exclusively on Shariff's example "b," which deploys a strictly asymmetrical bicolumnar grid. In this specific *mkondo*, the line is cleaved into two unequal hemistiches: the first hemistich is locked at exactly 4 syllables ($N = 4$), and the second hemistich is locked at exactly 8 syllables ($M = 8$).

Formal Parameters:

The poem is a tercet ($n_{max} = 3$).

The line is divided into two hemistiches ($H = 2$).

The total syllable count is 12 syllables per line, distributed as $S = 4 + 8$.

The Matrix:

L_1: [4, a] + [8, b] //
L_2: [4, a] + [8, b] //
L_3: [4, a] + [8, b] //

The Algebraic Equation for the Asymmetrical *Wimbo*:

$$\boldsymbol{Wimbo_{asymmetrical} = 3(4_a + 8_b//)}$$

3. Verification and the Availability of the Tetrasyllabic Template

This configuration is of particular theoretical importance, because it is the ground most favorable to the exogenous accentual school. Minerba leverages this type of structure to support his trochaic

innovation (Minerba, 2025, pp. 217–219), and he states the case for the four-syllable hemistich without qualification:

> [...] the four syllable hemistich of [*sic*] can contain either a phonetic word of 4 syllables (such as, for example *amekuja*, *na wazazi*) or two words of two syllables each (*watu wote*): the stresses are therefore always on odd syllables and the trochaic rhythm is necessarily reproduced. (Minerba, 2025, p. 221)

The enumeration is exhaustive and the conclusion follows, as section 2.5.2 concedes. A segment of four syllables admits only these two partitions once monosyllabic formatives are treated as proclitic, and both place the stress on an odd position. The first hemistich of this matrix will therefore satisfy the parity criterion in every line, and no measurement can show otherwise.

What this establishes is narrower than what it is invoked to support, and the case study is constructed to make the difference visible. The stress falls where it does because a four-syllable segment leaves the penultimate rule no other position to occupy, not because a tetrasyllabic template has been imported and applied. The question the verification can settle is therefore not whether the short hemistich is trochaic — it cannot be otherwise — but whether the eight-syllable hemistich, where the language does leave a choice, exhibits the same rhythm. Condition 6 governs that question, in both its strict form and its extended parity form; Condition 8 is not decisive here, since 4 and 8 are both multiples of four and the template is arithmetically available throughout.

The verification follows the three-step protocol of section 9.2.3. The accentual structure of every line is first derived from the printed text by the penultimate rule, an operation requiring no recording. The formal reading is then screened by ear, and the two results are reported side by side. Acoustic extraction is applied to a sample of lines only, and no statistic in this case study rests upon it. The reading maintains fidelity to the macro-Swahili phonology that governs this *wimbo*, whatever the northern coastal environment from which Shariff anthologized it.

4. Wimbo 4 + 8: Syllabic and Accentual Mapping

4.a. Reading the entries

Each line occupies a block of three short lines, in the notation of sections QRG.1 and 2.5. The first gives the reference and the line as printed, syllabified, with the hemistiches separated by the caesura marker + and their *mizani* counts in brackets. The second is the lexical reading, derived from the printed text by the penultimate rule: for each hemistich, the sequence of nucleus volumes, the positional index P*i* of each nucleus, and G-lex. Positions reset at each caesura. The third is the auditory reading, followed by Δ.

A question mark marks a weak realization: a nucleus counted although its prominence is markedly lower than that of the others in the same hemistich. The mark records a property of the reading, not a hesitation of the listener.

4.b. Mapping

Stanza 1 — rhyme a = *li*, rhyme b = *ne*
1.1 Si-na ha-li + mo-yo wa-ngu ni-de-nge-ne (4+8)
lexical 2·2 + 2·2·4 — P_1=1, P_2=3 | P_1=1, P_2=3, P_3=7 — G-lex 2 + 3
auditory 2?·2 + 2?·2·4 — P_1=1?, P_2=3 | P_1=1?, P_2=3, P_3=7
— G-aud 2 + 3 — Δ — | —
1.2 Si-hi-mi-li + wa-la si-pe-ndi m-ngi-ne (4+8)
lexical 4 + 2·3·3 — P_1=3 | P_1=1, P_2=4, P_3=7 — G-lex 1 + 3
auditory 4 + 2?·3·3 — P_1=3 | P_1=1?, P_2=4, P_3=7 — G-aud 1 + 3 — Δ — | —
1.3 Ta-fa-dha-li + na-ku-o-mba tu-o-na-ne (4+8)
lexical 4 + 4·4 — P_1=3 | P_1=3, P_2=7 — G-lex 1 + 2
auditory 4 + 4·4 — P_1=3 | P_1=3, P_2=7 — G-aud 1 + 2 — Δ — | —

4.c. Summary

The fragment comprises three lines, that is six hemistiches: three of four syllables and three of eight. No line is excluded, and the syllabic constant 4 + 8 holds in all three.

The four-syllable hemistiches. The three exhibit the two partitions that a segment of four syllables admits once monosyllabic formatives are treated as proclitic: 2·2 at 1.1, with stresses on positions 1 and 3, and a single nucleus of four at 1.2 and 1.3, with the stress on position 3. Both place every stress on an odd position. No other outcome was available, and the reading produces none.

The eight-syllable hemistiches. The lexical sequences are 2·2·4, 2·3·3 and 4·4, and the reading realizes each of them exactly: the departure rate is zero in three hemistiches. The sequence 4·4 occurs once in three. Three of the six hemistiches carry a nucleus marked as weakly realized, all of them the first nucleus of their hemistich.

5. Statistical Analysis and Falsification

5.a. What is tested here, and what is not

This corpus is a fragment of three lines, and the limits of what three lines can establish were set out in section 1. The conditions of section 9.2.4 are applied to it, but their weight differs from segment to segment, and the difference must be stated before any figure is given.

The 4 + 8 matrix offers two containers of unequal length, and they do not answer the same question.

The three four-syllable hemistiches are not tested for accentual structure. A segment of four syllables admits exactly two partitions once monosyllabic formatives are treated as proclitic: a single nucleus of four, or two nuclei of two. Both place every stress on an odd position. Minerba states this without qualification (Minerba, 2025, p. 221) and the statement is correct, as section 2.5.2 concedes. No reading can contradict it, and counting these three hemistiches as confirmations of the trochaic rhythm would be counting an arithmetical necessity as evidence. They are reported below for what they show about the composer, not about the meter.

The three eight-syllable hemistiches are tested as *utenzi* lines are, since they are containers of eight syllables filled by the Swahili lexicon under a metrical constraint. Three hemistiches support no proportion worth reporting as a percentage, and the figures below are given as counts.

Condition 3 is applicable, the line having an internal caesura. **Condition 4 is not**, the fragment employing no permutation or chaining.

5.b. Conditions 1 and 2

Condition 1. The matrix requires 4 + 8 in every line. The constant holds in all three, without deviation.

Condition 2. The local rhyme *a* binds the three four-syllable hemistiches on *li — hali*, *sihimili*, *tafadhali* — and the rhyme *b* binds the three eight-syllable hemistiches on *ne — nidengene*, *mngine*,

tuonane. Both hold in their syllabic form, a^s, and the vocalic form is not invoked.

Condition 3. No word crosses the caesura in any of the three lines. The boundary after the fourth syllable coincides with a word boundary throughout.

5.c. Condition 5: the bipartite axiom

On the three eight-syllable hemistiches, the lexicon yields two accentual groups in **one** and three in **two**. The reading realizes the same distribution: the departure rate is **zero in three hemistiches**, every sequence being realized exactly as the lexicon supplies it.

Verdict. The axiom is contradicted in two hemistiches of three. On so small a fragment no proportion is meaningful, and the result is recorded rather than interpreted.

5.d. Condition 6: the *Grenzschema*

The sequence 4·4 occurs in **one** of the three eight-syllable hemistiches, at 1.3, and is realized there. The two others carry 2·2·4 and 2·3·3, both of three nuclei of unequal volume.

The four-syllable hemistiches confirm Minerba, and the confirmation is worth examining rather than counting. The three exhibit both available partitions: 2·2 at 1.1, with stresses on positions 1 and 3, and a single nucleus of four at 1.2 and 1.3, with the stress on position 3. Every stress falls on an odd position, in all three lines, and no other outcome was possible.

What this establishes is narrower than what it is invoked to support. The stress falls where it does because a four-syllable segment leaves the penultimate rule no other position to occupy, not because a tetrasyllabic template has been imported and applied. The same configuration would arise in any four-syllable stretch of Swahili prose, and the constancy of the result across the three lines measures the rigidity of the constraint rather than the presence of a rhythm.

The comparison between the two hemistich lengths is the substance of this case study. Where the segment leaves the language no choice, the trochaic rhythm appears in three lines of three. Where the segment leaves a choice — the eight-syllable hemistich, with its nine available sequences — it appears in one line of three.

5.e. Condition 7: periodicity

The nuclei of a hemistich are all of equal volume in **one** of the three eight-syllable hemistiches, the 4·4 at 1.3, and in **one** of the three four-syllable hemistiches, the 2·2 at 1.1. In the four others no unit repeats.

Across hemistiches, three lines offer two consecutive pairs and no sequence recurs. The fragment is too short for the independence test of Condition 7 to be applied, and it is not applied.

5.f. Condition 8: the tetrasyllabic foot

Every segment of this matrix is a multiple of four, since 4 and 8 both are. **Condition 8 cannot be falsified on this corpus**, and this is precisely why the matrix was selected: it is the configuration most favorable to the imported template, the ground the accentual school would choose.

Where the template is available, it is used in **two of the six hemistiches**: the single nucleus of four at 1.2 and 1.3 — where, again, no alternative existed — and the 4·4 at 1.3. In the eight-syllable hemistiches of 1.1 and 1.2 it does not occur at all.

Verdict. On the ground most favorable to it, the template governs one eight-syllable hemistich of three. The condition is not decisive here and is tested where it can be, on the hyper-fractionated matrix of section 9.5.3.

9.5.3. Case Study V: *Wimbo* — Test of a Hyper-Fractionated Matrix — Untitled, by Sheikh Ahmed Sheikh Nabhany (1927–2017)

To complete the systematic testing of the exogenous accentual models, the extreme boundaries of the classical Swahili metrical grid must be brought into the corpus. In *Tungo Zetu* (1988, pp. 47–48), Ibrahim Noor Shariff provides a taxonomy of the *wimbo* demonstrating that the form is not constrained by a simple repeating rhythm but is capable of considerable geometric complexity. The furthest extension of this spatial manipulation is what we term the hyper-fractionated *wimbo*.

1. Presentation, Dating, and Philological Context

The text selected for this experimental verification is a *wimbo* composed by the Swahili master poet and scholar, Sheikh Ahmed Sheikh Nabhany (1927–2017). Documented by Shariff as example "k"

in his analysis of *wimbo* sub-meters (*mikondo*), this composition serves as a definitive testament to the unbroken continuity and architectural vitality of the classical grid well into the contemporary era.

Shariff introduces this text to explicitly demonstrate that a *wimbo* can harbor more than three rhymes per line, provided it respects the overarching three-line stanzaic limit ($n_{max} = 3$). He praises Nabhany as a *bingwa* (master) for engineering a highly intricate *mkondo* (sub-meter) structured upon four distinct hemistiches per line, creating a cascading sequence of rhymes. This is not a folk song driven by a drumbeat; it is a arithmetically rigorous, hyper-fractionated algebraic equation spatialized on the page.

2. The Hyper-Fractionated Matrix

The architectural blueprint of Nabhany's composition shatters any notion of a simple bipartite line. The verse is cleaved into four independent segments ($H = 4$), each governed by its own absolute syllable count and locked by an independent, interlocking rhyme chain (a, b, c, d). The syllabic sequence established by the poet is strictly $6 + 3 + 4 + 2$.

Formal Parameters for the 4-Hemistich *Wimbo*:

The poem is a tercet ($n_{max} = 3$).

The line is divided into four hemistiches ($H = 4$).

The total syllable count is 15 syllables per line. It is distributed as $S = 6 + 3 + 4 + 2$.

The Matrix:

L_1: [6, a] + [3, b] + [4, c] + [2, d] //
L_2: [6, a] + [3, b] + [4, c] + [2, d] //
L_3: [6, a] + [3, b] + [4, c] + [2, d] //

The Algebraic Equation for the Hyper-Fractionated *Wimbo*:

$$\boldsymbol{Wimbo_{fractionated} = 3(6_a + 3_b + 4_c + 2_d//)}$$

This matrix is the configuration on which Condition 8 becomes decisive, and the reason is arithmetical rather than statistical. Minerba derives the classical Swahili line from a tetrasyllabic template and states the derivation for the three meters examined in this study (Minerba, 2025, p. 217); the formalization of that claim and its conditions of failure are set out in section 2.5.2. A segment of three syllables cannot host a four-syllable foot, and a segment of two cannot host one either. The sequence 6 + 3 + 4 + 2 admits no partition into tetrasyllabic units by any arrangement whatever: of its four segments, one is a multiple of four, one is a multiple of three, and two are neither. The model does not predict badly here. It produces no prediction at all, while continuing to be asserted of the meter that contains this line.

3. Verification

The verification follows the three-step protocol of section 9.2.3, and the order of the steps matters to what each establishes.

The accentual structure of every line is first derived from the printed text by the penultimate rule. This operation requires no recording, no listener and no software, and any reader who knows the rule recomputes it from the text as printed. The formal reading is then screened by ear, and the two results are reported side by side. Acoustic extraction is applied to a sample of lines only, and no statistic in this case study rests upon it.

It is epistemologically vital to restate the distinction on which this protocol rests (see 3.3, *The 3-Layer Theory of Swahili Poetics*). The recording captures the text being formally read, not sung, which sets aside Layer 3, the melodic overlay of the performer.

What the recording gives access to is not a layer of the metrical system at all. Layers 1 and 2 — the syllabic law and the architecture built upon it — say nothing about accent. Accent belongs to the language: Swahili assigns one stress to each phonological word, on its penultimate syllable, so that a given sequence of words yields one accentual structure and only one. This structure is a property of the words the composer selected, and it is computable from the printed line without any recording.

What the reader receives is the page itself, in ordinary Swahili orthography: the words of the poem, laid out line by line, whose syllable counts and rhyme boundaries constitute Layers 1 and 2. Nothing is marked on it. There is no accentual notation, no indication

of syllable count, no typographical highlighting of the rhyming syllables, and no segmentation of the line into hemistiches.

The reading is what allows us to observe how much of the lexical structure a speaker actually realizes, since a formal reading superimposes a phrase intonation under which some nuclei are realized and others are not. Two datasets are therefore kept apart throughout: the accentual structure the lexicon supplies, and its partial realization in the reading. It is precisely the confusion of these two, and of both with the melodic freedom of Layer 3, that has sustained the accentual reading of Swahili meter for a century.

4. Hyper-fractionated Wimbo: Syllabic and Accentual Mapping

4.a. Reading the entries

Each line occupies a block of three short lines, in the notation of sections QRG.1 and 2.5. The first gives the reference and the line as printed, syllabified, with the hemistiches separated by the caesura marker + and their syllable counts *(mizani)* in brackets. The second is the lexical reading, derived from the printed text by the penultimate rule: for each hemistich, the sequence of nucleus volumes, the positional index P*i* of each nucleus, and G-lex. Positions reset at each caesura. The third is the auditory reading, followed by Δ.

A question mark marks a weak realization: a nucleus counted although its prominence is markedly lower than that of the others in the same hemistich. The mark records a property of the reading, not a hesitation of the listener.

4.b. Mapping

Stanza 1 — rhyme a = *na*, rhyme b = *ki*, rhyme c = *na*, rhyme d = *ni*
1.1 Ku-kwe-pu-ka na-na + si-ta-ki + na-so-no-na + nda-ni (6+3+4+2)
lexical 4·2 + 3 + 4 + 2 — P_1=3, P_2=5 | P_1=2 | P_1=3 | P_1=1
— G-lex 2 + 1 + 1 + 1
excluded — *guni* of reading: the reader executes *ku-kwe-pa* for *ku-kwe-pu-ka*, yielding five syllables where the source carries six

1.2 Ku-su-bi-ri te-na + ni dhi-ki + mu-hi-ba-na + li-ni (6+3+4+2)
lexical 4·2 + 3 + 4 + 2 — P1=3, P2=5 | P1=2 | P1=3 | P1=1
— G-lex 2 + 1 + 1 + 1
auditory 4·2 + 3 + 4 + 2 — P1=3, P2=5 | P1=2 | P1=3 | P1=1
— G-aud 2 + 1 + 1 + 1 — Δ — | — | — | —
1.3 Nyo-nda ku-o-na-na + a-shi-ki + ha-li si-na + sha-ni (6+3+4+2)
lexical 2·4 + 3 + 2·2 + 2 — P1=1, P2=5 | P1=2 | P1=1, P2=3 | P1=1
— G-lex 2 + 1 + 2 + 1
auditory 2·4 + 3 + 2·2 + 2 — P1=1, P2=5 | P1=2 | P1=1, P2=3 | P1=1
— G-aud 2 + 1 + 2 + 1 — Δ — | — | — | —

4.c. Excluded lines

One line of the three is removed from the statistical pool. At 1.1 the reader executes *ku-kwe-pa* where the source carries *ku-kwe-pu-ka*, so that the opening hemistich is delivered with five syllables instead of six. The deviation belongs to the reading and not to the text, and the line is excluded entire.

4.d. Summary

The pool comprises **two lines of the three mapped**, that is **eight hemistiches**: two of six syllables, two of three, two of four and two of two. The syllabic sequence 6 + 3 + 4 + 2 holds without deviation in the two retained lines.

The reading departs from the lexical structure in no hemistich whatever. All eight are realized exactly as the lexicon supplies them, which is the lowest departure rate recorded anywhere in this chapter.

The four-syllable hemistiches. The two exhibit the two partitions available to a segment of that length: a single nucleus of four at 1.2, with the stress on position 3, and 2·2 at 1.3, with stresses on positions 1 and 3. Both place every stress on an odd position, as no other outcome was available.

The remaining six hemistiches — two of six syllables, two of three, two of two — carry the sequences 4·2, 2·4, 3 and 2. None of them admits a partition into tetrasyllabic units, and the consequence for Condition 8 is drawn in section 5 below.

5. Statistical Analysis and Falsification

5.a. What is tested here, and what is not

This corpus consists of two retained lines, and it is not examined for the same reason as the others. Its value is not statistical. A matrix of four hemistiches in the sequence 6 + 3 + 4 + 2 settles Condition 8 by arithmetic, and the settlement holds whatever the reading does; the eight hemistiches are then reported for what they show about the syllabic model, which continues to describe where the other ceases to compute.

The conditions therefore apply unevenly, and the unevenness is stated before the figures.

Conditions 5, 6 and 7 are not applied to most of this corpus. They concern the accentual structure of an eight-syllable container, and this matrix contains none: its segments measure six, three, four and two syllables. Applying the bipartite axiom to a segment of three, or the *Grenzschema* to a segment of two, would be to ask a question the container cannot answer. Where the segment length permits a meaningful observation — the two four-syllable hemistiches — the observation is made and its limits declared.

Condition 8 is applied in full, and it is decisive. This is the configuration for which the condition was formulated.

Conditions 1, 2 and 3 are applied to the line and to the stanza, as elsewhere. Condition 4 is not applicable.

5.b. Conditions 1, 2 and 3

Condition 1. The matrix requires 6 + 3 + 4 + 2 in every line. The sequence holds without deviation in the two retained lines, across all eight hemistiches.

Condition 2. Four independent rhyme chains bind the four hemistich positions: *a* on *na*, *b* on *ki*, *c* on *na*, *d* on *ni*. All four hold in their syllabic form, a^{s}, and the vocalic form is not invoked in any of them.

The result deserves to be stated for what it demands of the composer. A single line of fifteen syllables carries four distinct rhymes, each recurring at its own position across the stanza, and the shortest of the four segments consists of two syllables which must simultaneously complete the line and carry the *d* rhyme. Shariff introduces this text precisely to show that a *wimbo* may harbor more

than three rhymes per line, and he calls Nabhany a *bingwa* for engineering it.

Condition 3. No word crosses any of the three internal caesuras, in either retained line. Six boundaries, six word boundaries.

5.c. The reading

The reading departs from the lexical structure in **no hemistich whatever**. All eight are realized exactly as the lexicon supplies them, which is the lowest departure rate recorded anywhere in this chapter — zero against 16.7 percent in the *Wimbo wa Miti*, 12.5 percent in the prose control corpus, and 11 to 12 percent in the two *utenzi*.

The figure rests on eight hemistiches and no statistical claim is made on it. What can be said is that the segments of this matrix are short, that each is closed by a rhyme, and that the reader realizes the structure the lexicon supplies without a single alteration.

One observation is recorded here, and it belongs to this configuration alone. The caesuras of this matrix are not marked by silence: no pause separates the four segments, and the line is delivered without interruption. They are nonetheless audible, and by a different means. The nuclei adjacent to a caesura, on either side of it as far as the ear can judge, are realized with a prominence markedly greater than their position in the lexical structure would lead one to expect, and several of them would plausibly have been weaker had no metrical boundary stood there. The frontier has no temporal reality in this reading; it has a prosodic one. The observation is made by ear and is not measured; it is recorded as a hypothesis for a corpus large enough to test it.

Two remarks follow. This effect is not observed in the other *wimbo* corpora of this chapter, where the segments are longer and the caesuras fewer; it appears here, where a fifteen-syllable line carries three internal boundaries and no segment exceeds six syllables. And it bears on the accentual model rather than on the syllabic one: a prominence produced by a metrical boundary is a prominence the lexicon does not predict, and it originates in the syllabic architecture itself rather than in any rhythmic template. Where such an effect operates, an analyst who reads the beats off the realization is recording the consequence of the *mizani*, not the cause of it.

5.d. Condition 8: the tetrasyllabic foot

Minerba derives the classical Swahili line from an imported tetrasyllabic template and states the derivation for the *utenzi*, the *shairi* and the *wimbo* alike (Minerba, 2025, p. 217). This matrix is where the derivation can be tested without recourse to any measurement.

Of the **eight hemistiches** of the retained corpus, **two measure four syllables** and can host a tetrasyllabic foot. The six others cannot, and no partition rescues them: a segment of six admits 4 + 2 or 2 + 4 but not 4 + 4; a segment of three admits no four-syllable unit; a segment of two admits none either. **Six hemistiches of eight, that is 75 percent, lie outside the reach of the template**, and of the thirty syllables of a line, **twenty-two are contained in segments the template cannot describe**.

Verdict. The model does not predict badly on this corpus. It produces no prediction at all for three quarters of its metrical units, while continuing to be asserted of the meter that contains them.

And the two segments that can host the template confirm it, as they must. At 1.2 the four-syllable hemistich is a single nucleus of four with the stress on position 3; at 1.3 it is 2·2 with stresses on positions 1 and 3. Both place every stress on an odd position, because no other outcome was available. The corpus therefore delivers the two results the model predicts in the two places where it cannot fail, and none of the six where it cannot speak.

5.e. What the syllabic model accounts for on the same corpus

The comparison is worth making explicit, since the two models are being applied to the same eight hemistiches.

Where the accentual model produces no prediction for six of them, the syllabic matrix predicts all eight: the sequence 6 + 3 + 4 + 2 is verified without deviation in the two retained lines, and the four rhyme chains hold in their strict syllabic form at every position. The segments that make the tetrasyllabic template incalculable are precisely the segments the syllabic matrix specifies, names and locks with a rhyme.

The point is not that one model performs better than the other on a shared measure. It is that the configurations which defeat the imported template are the ones the endogenous description was built to describe, and the ones Shariff singles out as the achievement of a master.

9.5.4. Instrumental Verification of the Hostile Sample

Instrumental analysis is applied to a sample of hemistiches and to no more than that, as section 9.2.3 sets out. The exhaustive verification of these three corpora is auditory, and no statistic in this case study rests on the measurements reported below. What the sample establishes is that the categories counted in section 5 have measurable physical correlates, and where they do not, that too is reported.

1. Why the sample is reasoned and not drawn

In the *utenzi* corpora the two lines added to the reasoned categories were obtained by a random draw on a population of well over a hundred. No such draw is made here, and the reason should be stated rather than passed over.

The three *wimbo* corpora yield thirty-three retained hemistiches of eight syllables, of which thirty belong to the *Wimbo wa Miti* and three to the 4 + 8 fragment; the hyper-fractionated *wimbo* contains none at all, its segments measuring six, three, four and two syllables. A draw on that population would return a hemistich of the same poem four times in five, and would leave the two other corpora unrepresented in more than half of its outcomes. A procedure that guarantees impartiality on a hundred and forty lines guarantees nothing on thirty-three unevenly distributed ones, and applying it here would lend the appearance of a safeguard without its substance. The sample is therefore composed by category alone, and the categories are stated before the measurements.

2. Why the short hemistiches are excluded

The sample is restricted to hemistiches of eight syllables, and to one of six. The segments of two and four syllables are set aside, and the ground for setting them aside is arithmetical rather than practical.

A segment of two syllables admits one accentual nucleus and no other, the penultimate rule having no second position to occupy. The sixteen two-syllable hemistiches of the *Wimbo wa Miti* are realized with that single nucleus without exception, as section 4 reports. A segment of four syllables admits two partitions once monosyllabic formatives are treated as proclitic — a single nucleus of four, or two of two — and both place every stress on an odd position, as Minerba

observes (2025, p. 221) and as section 5 concedes. The five four-syllable hemistiches of these corpora exhibit both partitions and nothing else.

In neither case can a measurement contradict the reading, because no other reading was available to the language. Measuring these segments would produce figures that confirm an arithmetical necessity, and confirmation of a necessity is not evidence. They are therefore excluded from the sample, and the exclusion is declared here so that their absence is not mistaken for an oversight.

3. Composition of the sample

Five hemistiches are measured, and the composition was fixed before any measurement was made.

A hemistich conforming exactly to the *Grenzschema*. *Wimbo wa Miti*, stanza 8, line 2, second hemistich, *nimeziye kutambuwa*: lexical sequence 4·4, stresses on positions 3 and 7, realized exactly. This is the accentual model's ideal matrix, measured first so that the model is tested where it is strongest.

A hemistich carrying the dominant lexical sequence of the corpus. *Wimbo wa Miti*, stanza 6, line 1, second hemistich, *hatuchi twapo uwawa*: sequence 3·2·3, stresses on positions 2, 4 and 7, realized exactly. This sequence, and not 4·4, is the most frequent in the corpus; it carries three nuclei of unequal volume and falls outside the trochaic rhythm on Minerba's own criterion.

A hemistich carrying three nuclei of decreasing volume. *Wimbo wa Miti*, stanza 5, line 1, second hemistich, *wasiyesa kani zao*: sequence 4·2·2, stresses on positions 3, 5 and 7, realized exactly. The configuration places three prominences in eight syllables at intervals of two, which is the densest accentual structure of the corpus and the furthest from the bipartite axiom.

A hemistich whose two nuclei are marked as weakly realized. *Wimbo wa Miti*, stanza 2, line 3, second hemistich, *kunyamaa ni ujura*: sequence 4·4 on positions 3 and 7, both nuclei heard but both markedly less prominent than elsewhere in the line. The confidence mark asserts a property of the signal, and the instrument reports on it.

A six-syllable segment of the hyper-fractionated matrix. *Hyper-fractionated Wimbo*, stanza 1, line 2, first hemistich, *Kusubiri tena*: sequence 4·2, stresses on positions 3 and 5, realized exactly. This segment admits no partition into tetrasyllabic units, and it is measured to show what the reading does with a length the imported template cannot describe.

4. What is measured and what is reported

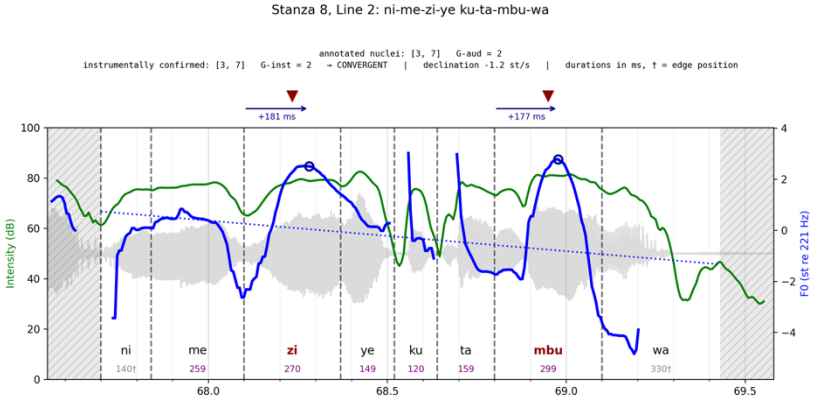
For each hemistich of the sample, three parameters are measured syllable by syllable: fundamental frequency in semitones relative to the speaker's median and corrected for the declination internal to the segment, syllable duration, and mean intensity. The analysis range is derived from the speaker's own distribution rather than set by hand, and the conditioning applied to the F0 track is stated in the header of every script.

Each figure carries three readings of the same hemistich. The lexical reading is derived from the printed text by the penultimate rule. The auditory reading is the annotation of section 4. The instrumental reading is established without reference to either: a syllable counts as instrumentally prominent when its detrended F0 mean is a local maximum across the syllables of the segment and lies above the declination. That criterion is free to confirm a nucleus the lexicon predicts, to leave one unconfirmed, or to designate a syllable that neither the lexicon nor the ear proposes.

Where the readings diverge, the divergence is stated in the caption, position by position, and the auditory annotation is not revised. Fundamental frequency is one parameter of three and is not decisive on its own; the durations and intensities are printed beneath every figure whether or not they agree with it.

5. Figures

Figure 1. *Wimbo wa Miti*, stanza 8, line 2, second hemistich: *nimeziye kutambuwa*.

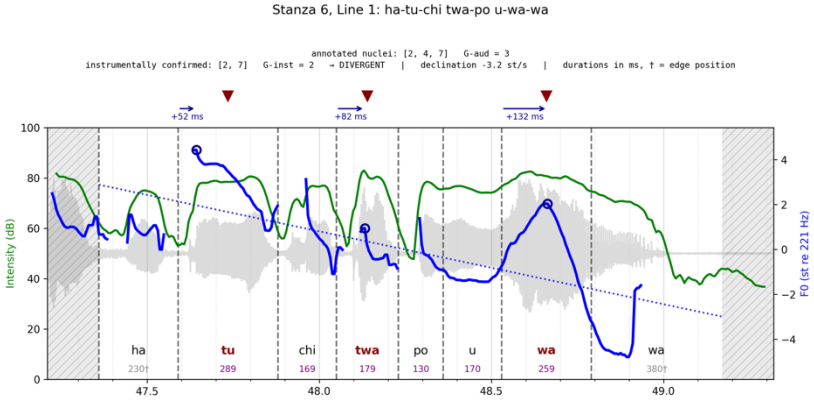



```
Wimbo wa Miti, Stanza 8, Line 2, H2: "ni-me-zi-ye ku-ta-mbu-wa"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4·4, P₁=3, P₂=7, G-aud = 2, delta —.
Instrumental reading: 4·4, P₁=3, P₂=7, G-inst = 2, delta — (against
the lexical reading).
Speaker median 221 Hz; declination −1.20 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 ni     140 ms†  F0 -1.15 st  peak +133 ms  intensity  71.3 dB
  2 me     259 ms   F0 -0.46 st  peak + 88 ms  intensity  75.5 dB
  3 zi     270 ms   F0 +0.91 st  peak +181 ms  intensity  75.7 dB
  4 ye     149 ms   F0 +0.69 st  peak +  0 ms  intensity  75.0 dB
  5 ku     120 ms   F0 +0.14 st  peak + 39 ms  intensity  62.9 dB
  6 ta     159 ms   F0 -0.23 st  peak + 54 ms  intensity  70.5 dB
  7 mbu    299 ms   F0 +0.87 st  peak +177 ms  intensity  77.8 dB
  8 wa     330 ms†  F0 -3.19 st  peak +  1 ms  intensity  57.0 dB
```

Figure 2. *Wimbo wa Miti*, stanza 6, line 1, second hemistich: *hatuchi twapo uwawa*.



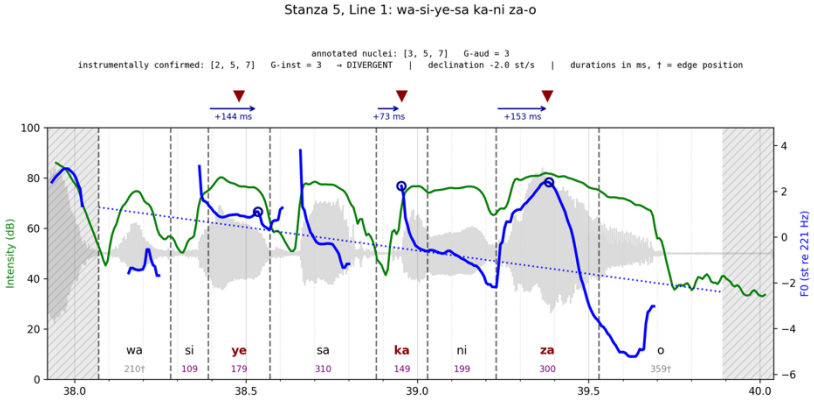
Wimbo wa Miti, Stanza 6, Line 1, H2: "ha-tu-chi twa-po u-wa-wa"
Lexical reading: 3·2·3, P_1=2, P_2=4, P_3=7, G-lex = 3.
Auditory reading: 3·2·3, P_1=2, P_2=4, P_3=7, G-aud = 3, delta –.
Instrumental reading: 3·5, P_1=2, P_2=7, G-inst = 2, delta $-P_2$=4 (against the lexical reading).
Speaker median 221 Hz; declination –3.25 st/s.

The instrumental reading diverges from the auditory reading: position(s) [4] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	ha	230 ms†	F0 -1.75 st	peak + 94 ms	intensity	64.4 dB
2	tu	289 ms	F0 +1.25 st	peak + 52 ms	intensity	73.8 dB
3	chi	169 ms	F0 -0.34 st	peak + 81 ms	intensity	70.1 dB
4	twa	179 ms	F0 -0.46 st	peak + 82 ms	intensity	71.7 dB
5	po	130 ms	F0 -0.09 st	peak + 61 ms	intensity	66.1 dB
6	u	170 ms	F0 -0.65 st	peak +167 ms	intensity	75.4 dB
7	wa	259 ms	F0 +1.64 st	peak +132 ms	intensity	79.7 dB
8	wa	380 ms†	F0 -1.95 st	peak +144 ms	intensity	56.7 dB

Figure 3. *Wimbo wa Miti*, stanza 5, line 1, second hemistich: *wasiyesa kani zao*.



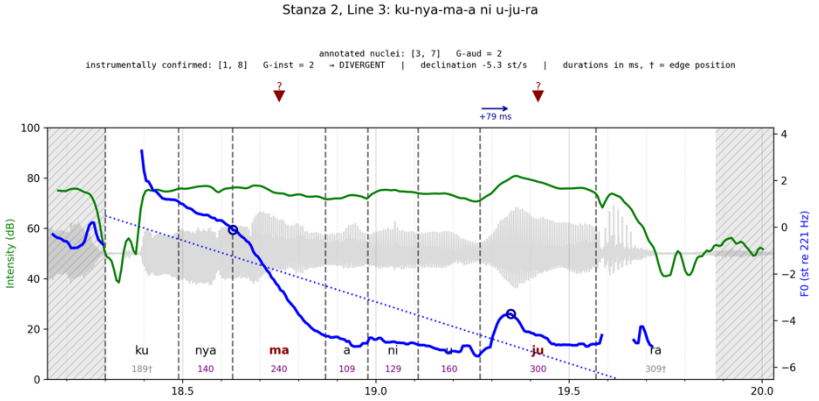
Wimbo wa Miti, Stanza 5, Line 1, H2: "wa-si-ye-sa ka-ni za-o"
Lexical reading: 4·2·2, P_1=3, P_2=5, P_3=7, G-lex = 3.
Auditory reading: 4·2·2, P_1=3, P_2=5, P_3=7, G-aud = 3, delta -.
Instrumental reading: 3·3·2, P_1=2, P_2=5, P_3=7, G-inst = 3, delta $-P_1$=3, $+P_1$=2 (against the lexical reading).
Speaker median 221 Hz; declination –2.03 st/s.

The instrumental reading diverges from the auditory reading: position(s) [3] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	wa	210 ms†	F0 -2.32 st	peak +146 ms	intensity	62.9 dB
2	si	109 ms	F0 +1.23 st	peak + 83 ms	intensity	59.7 dB
3	ye	179 ms	F0 +0.44 st	peak +144 ms	intensity	76.1 dB
4	sa	310 ms	F0 +0.07 st	peak + 88 ms	intensity	67.1 dB
5	ka	149 ms	F0 +0.75 st	peak + 73 ms	intensity	63.7 dB
6	ni	199 ms	F0 -0.24 st	peak + 64 ms	intensity	74.1 dB
7	za	300 ms	F0 +1.72 st	peak +153 ms	intensity	78.0 dB
8	o	359 ms†	F0 -2.63 st	peak +160 ms	intensity	54.7 dB

Figure 4. *Wimbo wa Miti*, stanza 2, line 3, second hemistich: *kunyamaa ni ujura*.

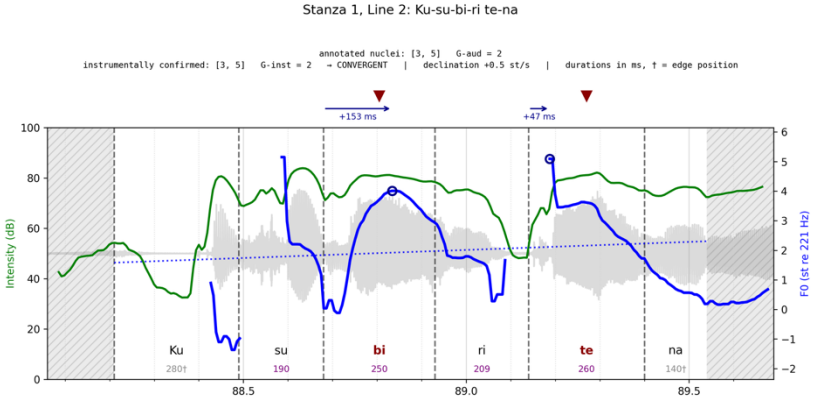



```
Wimbo wa Miti, Stanza 2, Line 3, H2: "ku-nya-ma-a ni u-ju-ra"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4?·4?, P₁=3?, P₂=7?, G-aud = 2, delta —.
Instrumental reading: 2·6, P₁=1, P₂=8, G-inst = 2, delta −P₁=3, −P₂=7,
+P₁=1, +P₂=8 (against the lexical reading).
Speaker median 221 Hz; declination −5.27 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [3, 7] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [1, 8] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 ku     189 ms†  F0 +1.83 st  peak + 94 ms  intensity  62.5 dB
  2 nya    140 ms   F0 +1.39 st  peak + 10 ms  intensity  75.9 dB
  3 ma     240 ms   F0 -0.59 st  peak +  0 ms  intensity  74.4 dB
  4 a      109 ms   F0 -2.08 st  peak +107 ms  intensity  72.3 dB
  5 ni     129 ms   F0 -1.46 st  peak +121 ms  intensity  74.6 dB
  6 u      160 ms   F0 -1.03 st  peak +133 ms  intensity  72.8 dB
  7 ju     300 ms   F0 +0.74 st  peak + 79 ms  intensity  76.8 dB
  8 ra     309 ms†  F0 +1.92 st  peak +121 ms  intensity  56.4 dB
```

Figure 5. ***Hyper-fractionated Wimbo*****, stanza 1, line 2, first hemistich:** ***Kusubiri tena*****.**




```
Hyper-fractionated Wimbo, Stanza 1, Line 2, H1: "Ku-su-bi-ri te-na"
Lexical reading: 4·2, P₁=3, P₂=5, G-lex = 2.
Auditory reading: 4·2, P₁=3, P₂=5, G-aud = 2, delta —.
Instrumental reading: 4·2, P₁=3, P₂=5, G-inst = 2, delta — (vs G-lex)
Speaker median 221 Hz; declination +0.54 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 Ku     280 ms†  F0 -2.46 st  peak +216 ms  intensity  51.8 dB
  2 su     190 ms   F0 +0.42 st  peak + 96 ms  intensity  76.1 dB
  3 bi     250 ms   F0 +0.75 st  peak +153 ms  intensity  78.3 dB
  4 ri     209 ms   F0 -0.44 st  peak +  3 ms  intensity  66.7 dB
  5 te     260 ms   F0 +0.90 st  peak + 47 ms  intensity  76.0 dB
  6 na     140 ms†  F0 -1.46 st  peak +  5 ms  intensity  74.7 dB
```

6. The concordance of the instrumental reading

The protocol of section 9.2.3 requires that this step report a concordance rate together with its failures. Of the five hemistiches measured, the instrumental reading coincides exactly with the auditory annotation in **two** — the *Grenzschema* at 8.2 and the six-syllable segment of the hyper-fractionated matrix — and departs from it in **three**. The rate is therefore two of five, and the departures are set out below rather than summarized away.

The two convergences fall on the two segments where the language leaves least room. At 8.2 the lexicon supplies two words of four syllables and the three readings agree without reserve: lexical

4·4, auditory 4·4, instrumental 4·4, on positions 3 and 7. At the six-syllable segment of the hyper-fractionated matrix the lexicon supplies a four-syllable word followed by a two-syllable one, and again the three readings agree on 4·2, positions 3 and 5. Where the words determine the structure closely, ear and instrument report the same thing.

The confidence mark predicts the largest divergence of the sample. The hemistich at 2.3 is the only one whose two nuclei were both marked as weakly realized, and it is the only one where the instrumental reading bears no relation whatever to the auditory one: the ear places the prominences on positions 3 and 7, the instrument places them on positions **1 and 8**, the two edge syllables. This hemistich carries the steepest declination of the sample at −5.27 semitones per second, and once that trend is removed the highest detrended values fall on the line-initial reset, *ku* at +1.83, and on the final syllable, *ra* at +1.92, while the syllables the ear hears as prominent measure −0.59 and +0.74. What the instrument records here is not the accentual structure of the hemistich: it is the pitch reset that opens a segment and the fall that closes it, in a hemistich where the melodic movement is so steep that the lexical prominences are buried within it. The reader's own report — two nuclei heard but both weak — turns out to describe exactly the condition under which the melodic parameter ceases to be informative.

One divergence reduces G, and it reduces it toward the accentual model. At 6.1 the ear hears three nuclei, 3·2·3 on positions 2, 4 and 7, which is the dominant lexical sequence of this corpus. The instrument confirms only two, on positions 2 and 7, the middle nucleus *twa* measuring −0.46 semitones against +1.25 for *tu* and +1.64 for *wa*. Read instrumentally, the hemistich becomes 3·5 with two accentual groups, which satisfies the bipartite axiom that the ear and the lexicon both deny. The observation should be stated plainly: where the two procedures disagree here, the instrumental one returns the result more favorable to the model under test, and this study reports it as such.

One divergence is a displacement of a single syllable. At 5.1 the instrument places the first prominence on position 2 where the ear and the lexicon place it on 3, the syllable *si* measuring +1.23 against +0.44 for *ye*. The displacement runs earlier, not later, and the two other parameters side with the ear: ye is both the longer and the louder of the two syllables (179 milliseconds against 109, 76.1 dB against 59.7). The contour, read syllable by syllable, follows pitch alone; the ear weighs pitch, length and loudness together.

7. What this establishes about the two models.

For the syllabic model, nothing, and the point is the same as in the *utenzi* corpora. The syllable count and the rhyme boundaries are properties of the printed text, and no measurement of prominence can confirm or threaten them. That the instrumental reading varies from hemistich to hemistich leaves Conditions 1, 2 and 3 exactly where section 5 left them. It is worth noting, as an empirical corollary, that the syllabic segmentation is visually corroborated across the sample figures by the acoustic intensity track, plotted in green: the successive energetic peaks consistently align with the syllabic nuclei, demonstrating the physical reality of the absolute syllable count independently of the fluctuating Fo contour.

For the accentual model, the sample confirms on five hemistiches what the *utenzi* corpora established on sixteen. Two readings of the same performance, obtained by procedures the accentual tradition would accept, disagree in three cases of five, and in one of them they return prominences on opposite ends of the segment. The accentual tradition since Werner has located its beats by ear, most probably in performance, where the most salient pitch movements belong to the melody rather than to the words; read off the raw contour instead, as an instrument reads them, the same movements yield these displacements and boundary tones. Either way, acoustic surface phenomena are taken for generative structural laws.

8. Reproduction

One script corresponds to each hemistich of the sample. Each carries in its header the accentual analysis of the segment, its boundaries in the recording and the onset of every syllable, and each produces its figure and its caption from the audio alone. The scripts and the recording are deposited as described in section 9.2.7.

9.5.5. Conclusion on *Wimbo* Meter

Three corpora have been examined in this meter, under the same protocol and by the same reader: the *Wimbo wa Miti*, a Rupture Matrix alternating 2 + 8 with 8 + 8; the 4 + 8 fragment anthologized by Shariff; and Nabhany's hyper-fractionated matrix in 6 + 3 + 4 + 2. Thirty lines were mapped, two removed for reasons of execution, and **twenty-eight retained, comprising sixty hemistiches**. The

figures below are those of the three case studies, restated together, and each condition is closed with the verdict its figures support.

The *wimbo* is not tested for the same thing as the *utenzi*, and the difference governs everything that follows. An *utenzi* corpus offers one container of eight syllables, repeated. The *wimbo* offers containers of two, three, four, six and eight syllables within a single meter, and it is this variety, and not any statistical result, that makes the meter decisive for the model under test.

1. The syllabic model

Condition 1. Three different syllabic matrices were applied, and the constants hold in **all sixty hemistiches of the twenty-eight retained lines, without a single exception**. No *guni* of the source text occurs in any of the three corpora. Two lines of the *Wimbo wa Miti* show the constraint operating on the composer rather than describing her: at 5.2 a syllabic nasal supplies the second syllable the two-syllable hemistich requires, and at 6.3 the standard *kuamua*, four syllables, is contracted to *muwa*, two, so that the hemistich closes on exactly eight.

Condition 2. Every rhyme chain of the three corpora holds in its strict syllabic form, a^s, and the vocalic license is invoked nowhere. This includes the four independent chains of the hyper-fractionated matrix, one of which must be carried by a hemistich of two syllables, and the eight stanzas of the *Wimbo wa Miti*, where a two-syllable hemistich rhymes with an eight-syllable one on the same full syllable.

Condition 3, applicable for the first time in this chapter. No word crosses a caesura anywhere in the three corpora: **sixty-eight boundaries, sixty-eight word boundaries**. The result is at its most demanding in the hyper-fractionated matrix, where a fifteen-syllable line carries three internal boundaries, and in the *Wimbo wa Miti*, where sixteen boundaries fall after two syllables.

Condition 4 is not applicable, none of the three corpora employing permutation or stanzaic chaining.

Verdict. The syllabic matrices describe the three corpora exhaustively. Every hemistich has the length its position prescribes, every rhyme chain holds in its strongest form, and no word crosses a metrical boundary. There is nothing left over for another model to explain.

2. The accentual model

2.a. Where the model can be tested at all

Of the **sixty retained hemistiches, thirty-three measure eight syllables** — thirty in the *Wimbo wa Miti*, three in the 4 + 8 fragment. These are the only containers to which Conditions 5, 6 and 7 can be meaningfully applied, and the figures below are computed on them.

The remaining twenty-seven are of two, three, four and six syllables, and they are not tested for accentual structure. A segment of two syllables admits one nucleus and no other; a segment of four admits two partitions, both trochaic, as Minerba correctly observes (2025, p. 221). Counting these as outcomes would be counting an arithmetical necessity as evidence. They are examined instead under Condition 8, where their length is the point at issue.

2.b. Is the bipartite axiom verified?

On the thirty-three eight-syllable hemistiches, the lexicon yields exactly two accentual groups in **18, that is 54.5 percent**, three in 14 (42.4 percent) and four in one. The reading realizes two in **17 (51.5 percent)** and three in 16 (48.5 percent).

Verdict. The axiom is not verified. It is contradicted by 15 hemistiches on the lexical measure and by 16 on the auditory one, and on the auditory measure the two outcomes are within one hemistich of each other.

The transition. The reading departs from the lexical structure in **4 of the 33 hemistiches, that is 12.1 percent**, against **12.5 percent** in the prose control corpus of section 9.7. The two figures are the same to within a third of one hemistich. The hyper-fractionated corpus departs in no hemistich of eight, its short segments being realized exactly as the lexicon supplies them.

2.c. Is the *Grenzschema* a cognitive prototype?

The sequence 4·4 occurs in **8 hemistiches of 33, that is 24.2 percent**, on the lexical measure, and in **8, that is 24.2 percent**, on the auditory one. The reading neither adds an instance nor removes one.

Rank. The most frequent lexical sequence of this meter is not 4·4 but **3·2·3, with eight occurrences, level with it** — three nuclei of unequal volume, which Minerba classes outside the trochaic rhythm.

The extended parity criterion. Of the seventeen hemistiches realized with two accentual groups, 10 place the first stress on an odd syllable, that is 58.8 percent, and 7 do not.

Verdict, and it must be given in two parts. On the extended criterion the model holds for a majority of the two-group hemistiches, and this is the first meter of this chapter where it does; the figure is reported without qualification. On the strict criterion it governs fewer than one hemistich in four and is level in the lexicon with a sequence it classes outside the rhythm. Neither figure is what a generative prototype produces.

2.d. Is there a rhythm?

The nuclei of a hemistich are all of equal volume in **9 of the 33, that is 27.3 percent**. In the remaining 24 no unit repeats.

Across hemistiches, two consecutive eight-syllable segments of the *Wimbo wa Miti* carry the same sequence in **4 pairs of 29, that is 13.8 percent**, against 4.5 expected under independence, that is 15.6 percent — **below** what independent draws would produce.

Verdict. There is no rhythm at either scale. Three hemistiches in four contain no repeating unit, and the succession of configurations is not distinguishable from independent draws.

2.e. Condition 8, and why this meter is decisive

Minerba derives the classical Swahili line from an imported tetrasyllabic template and states the derivation for the *utenzi*, the *shairi* and the *wimbo* alike:

> The major innovation in the history of the *utenzi*, *shairi* and *wimbo* metres is their adoption of trochaic rhythms as constitutive elements of their abstract patterns. This is evident, historically, in their progressive adoption of syllabic lengths of multiples of four. (Minerba, 2025, p. 217)

Of the **sixty retained hemistiches of this meter, 22 have a length that is not a multiple of four**, that is **36.7 percent**: eighteen of two syllables, two of three, two of six. In these twenty-two segments a four-syllable foot cannot be constructed, and no partition rescues them. Measured in syllables, 54 of the 338 in the corpus lie in segments the template cannot describe.

Where the template is available, it is used in **12 of the 33 eight-syllable hemistiches, that is 36.4 percent**, and only 8 are partitioned entirely into tetrasyllabic units.

Verdict. For more than a third of the metrical units of this meter, the model does not predict badly: it produces no prediction at all, while continuing to be asserted of the meter that contains them.

2.f. What the twenty-two segments are

The verdict above would be a technical remark if the segments in question were marginal, and the accentual school's framework is built to make them appear so. Minerba classifies his measurements according to "their adherence to the theorized norm: fulfilment, exception, and error" (Minerba, 2025, p. 101); a line whose count departs from the expected value becomes an error (p. 102), a structurally inconvenient hemistich becomes an exception relying on prosodic contraction (p. 103). The asymmetrical configurations he himself lists — 6 + 6, 2 + 8, 6 + 4, 8 + 5 (p. 160) — are named and set aside. A framework that classifies every deviation from the theorized norm as a defect of the data cannot be contradicted by a deviation, and that is the definition of an unfalsifiable hypothesis.

The three corpora examined here permit that classification to be tested against what the segments actually are, and the answer is not in doubt.

They are not accidents of composition. The eighteen two-syllable hemistiches of the *Wimbo wa Miti* are not short by default: each occupies the head of a line, each is closed by the stanza's own rhyme, and the same shape recurs on the first and second line of all eight stanzas. The composer builds a boundary after two syllables, sixteen times, and locks it with a rhyme that must match an eight-syllable hemistich elsewhere in the stanza. She meets that constraint without a single license.

They are what the tradition calls mastery. Shariff introduces Nabhany's 6 + 3 + 4 + 2 precisely to show that a *wimbo* may carry more than three rhymes in a line, and he names Nabhany a *bingwa* for engineering it (Shariff, 1988, pp. 47–48). The configurations that make the imported template incalculable are the ones the endogenous description singles out as achievements. They are not the margin of the corpus; they are its most worked part.

And they are exactly what the syllabic matrix describes. The sixty hemistiches whose lengths defeat the tetrasyllabic derivation are the sixty the *mizani* specifies, positions and all, and

whose rhyme chains hold in their strict form at every boundary. Where one model ceases to compute, the other continues to predict, and what it predicts is the very constraint that made the first one stop.

2.g. Conclusion on this meter

The syllabic model accounts for 100 percent of the hemistiches of these three corpora on Conditions 1, 2 and 3, from the printed text, by a procedure any reader may repeat.

The accentual model accounts for 54.5 percent of the eight-syllable hemistiches on the count, 24.2 percent on the placement, and nothing at all on the 36.7 percent of hemistiches whose length forbids its unit. It supplies no repeating unit in three hemistiches of four and no sequencing between them beyond what independence produces. Its one favorable result — a bare majority on the extended parity criterion — is obtained on a subset of a subset, and it is reported here as favorable.

What separates the two models is not a difference of degree. The syllabic matrix was derived from the endogenous descriptions and confronted with texts those descriptions did not analyze, and no line contradicted it. The tetrasyllabic derivation was imported, extended to three meters in the space of five pages, and equipped in advance with a classification that converts every contradicting datum into an error or an exception. The first model is answerable to the corpus. The second has been arranged so that the corpus cannot answer it.

The same conditions are applied to the *shairi* in the case studies that follow, and the three partial results are drawn together in the conclusion of this book.

9.6. *Shairi* Texts: The Ultimate Complexity Tests

9.6.1. Case Study VI: Test of a Five-Line (6 + 9) Shairi — *Waji Waji (Dua ya Kuombea Mvua)*, by Sheikh Muhyddin bin Sheikh bin Jahathwan Al-Waily (1798–1870)

To rigorously test the explanatory power of the syllabic grid against the accentual hypothesis, we subject a 19th-century text to empirical analysis. For this verification, we utilize the poem *Waji Waji (Dua ya Kuombea Mvua),* a prayer for rain composed by Sheikh Muhyddin bin Sheikh bin Jahathwan Al-Waily (1798–1870). The author was an Islamic scholar and an expert in Swahili literature and other disciplines. He is also known for writing a history of Kilwa, which he presented to the Sultan of Zanzibar, who in turn gave it to the British ambassador; this manuscript was eventually taken to Britain and housed in a museum. Alongside *Waji Waji*, he also composed an *ukawafi* meter poem titled *Miraji*, recounting the story of the Prophet Mohammed's ascension to heaven. The poem analyzed here was published in the anthology *Malenga wa Karne Moja*, compiled by Shihabdin Chiraghdin with an introduction by Ahmed Sheikh Nabhany (Chiraghdin & Nabhany, 1987, pp. 45-47). The biographical overview of the author provided above is directly paraphrased from the prefatory notes found on these specific pages.

In order to confront the accentual school with the material evidence, the model tested here is the one formalized in section 2.5.2. Minerba derives the classical Swahili line from an imported tetrasyllabic template and states the derivation for the *utenzi*, the *shairi* and the *wimbo* alike (Minerba, 2025, p. 217). The model operates at two degrees of strictness, and both are tested: in its strict form it requires two tetrasyllabic nuclei with stresses on positions 3 and 7, which is the *Grenzschema*; in its extended form it requires only that the first stress fall on an odd syllable (Minerba, 2025, p. 220).

The syllabification and the accentual mapping follow the three-step protocol of section 9.2.3, restated in the preliminaries to this section. The accentual structure of each line is first derived from the printed text by the penultimate rule; the formal reading is then screened by ear, and the two results are reported side by side. Acoustic extraction

is applied to a sample of hemistiches only, and it is reported for the three *shairi* corpora together in section 9.6.4.

1. Identification of the Formal Syllabic Matrix

The structural architecture of this poem reveals a highly complex, yet perfectly regulated, asymmetrical configuration. The poem adheres to the foundational *shairi* rhyming principles, but it executes a radical spatial expansion. It breaks the classical four-line stanza (*tarbia*) into a five-line stanza (*takhmisa*, or *utano*) (see Chapter 4 for the Terminology). Even more devastating for exogenous rhythmic theories, it explicitly abandons the symmetrical 16-syllable line in favor of a 15-syllable line divided asymmetrically into 6 and 9 syllables. Furthermore, it utilizes the generative sub-meter known as ***Kupindua kina kimoja*** (the permutation of a single rhyme).

The formal generative matrix for one stanza (*ubeti*) operates as follows:

Formal Parameters:

$n_{max} = 5, H = 2, S = 15$ (where $S = 6 + 9$, therefore $N = 6$ and $M = 9$).

The Matrix:

L_1: $[6, a] + [9, b]$ / /
L_2: $[6, a] + [9, b]$ / /
L_3: $[6, a] + [9, b]$ / /
L_4: $[6, a] + [9, b]$ / /
L_5: $[6, b] + [9, X]$ / /

The Algebraic Equation for this specific stanzaic configuration:

$$\boldsymbol{Shairi}_{\boldsymbol{Takhmisa_|\ Kupindua\ kina\ kimoja_|\ Asymmetrical}} = \mathbf{4}([\mathbf{6}\boldsymbol{a}] + [\mathbf{9}\boldsymbol{b}]//) \cdot ([\mathbf{6}\boldsymbol{b}] + [\mathbf{9}\boldsymbol{X}]//)$$

Across the stanzas, the first four lines share an internal rhyme *a* and a terminal rhyme *b*, which act as local variables changing from one stanza to the next. In the fifth line, the preceding terminal rhyme *b* permutes to the position of the internal caesura, while the absolute terminal boundary is locked by the global constant rhyme *X* extending across the entire poem (which, in this composition, is the syllable -*ma*).

2. Waji Waji: Syllabic and Accentual Mapping, Stanzas 1–14

Each line is given as two hemistiches separated by the caesura marker +, with their syllable counts *(mizani)* in brackets. The notation, the derivation of the lexical reading by the penultimate rule and the value of the question mark are set out in the preliminaries to this section. Positions reset at the caesura. Where the lexical and the auditory readings coincide, a single line gives both; where they differ, the two are given separately and followed by Δ, which records the transition.

2.a. Mapping

Stanza 1 — rhyme a = *na*, rhyme b = *ka*
1.1 a-i su-bu-ha-na + mo-la we-tu shi-ri-ka (6+7)
lexical 2·4 + 2·2·3 — P_1=1, P_2=5 | P_1=1, P_2=3, P_3=6 — G-lex 2 + 3
auditory 2?·4 + 4·3 — P_1=1?, P_2=5 | P_1=3, P_2=6
— G-aud 2 + 2 — Δ — | −P_1=1 *guni of the text: seven syllables in H2*
1.2 ndi-we ra-ha-ma-na + mu-wa-wa-zi u-si-o sha-ka (6+9)
lexical = auditory 2·4 + 4·3?·2 — P_1=1, P_2=5 | P_1=3, P_2=6?, P_3=8 — G 2 + 3
1.3 wa-re-he-mu wa-na + na wa-ze-e we-nye ma-sha-ka (6+9)
lexical = auditory 4·2 + 4·2·3 — P_1=3, P_2=5 | P_1=3, P_2=5, P_3=8 — G 2 + 3
1.4 wa-i-ju-wa sa-na + ha-li ye-tu ka-ti cha-ka (6+8)
lexical 4·2 + 2·2·2·2 — P_1=3, P_2=5 | P_1=1, P_2=3, P_3=5, P_4=7 — G-lex 2 + 4
auditory 4·2 + 4·4 — P_1=3, P_2=5 | P_1=3, P_2=7
— G-aud 2 + 2 — Δ — | −P_1=1, −P_3=5 *guni of the text: eight syllables in H2*
1.5 ju-wa la m-ta-na + ku-la si-ku li-si-lo-ko-ma (6+9)
lexical 2·4 + 2·2·5 — P_1=1, P_2=5 | P_1=1, P_2=3, P_3=8 — G-lex 2 + 3
auditory 2·4 + 2·2·3·2 — P_1=1, P_2=5 | P_1=1, P_2=3, P_3=6, P_4=8
— G-aud 2 + 4 — Δ — | +P_3=6
Stanza 2 — rhyme a = *fu*, rhyme b = *le*
2.1 tu wa-ja dha-i-fu + m-la-ngo ni-pa-ko tu-e-me-le (6+10)
lexical = auditory 3·3 + 3·3·4 — P_1=2, P_2=5 | P_1=2, P_2=5, P_3=9 — G 2 + 3
guni of the text: ten syllables in H2
2.2 kwa ni-a ta-ka-fu + shi-da ye-tu ra-bi o-ndo-le (6+9)
lexical 3·3 + 2·2·2·3 — P_1=2, P_2=5 | P_1=1, P_2=3, P_3=5, P_4=8 — G-lex 2 + 4
auditory 2·4 + 2?·2·2·3 — P_1=1, P_2=5 | P_1=1?, P_2=3, P_3=5, P_4=8
— G-aud 2 + 4 — Δ −P_1=2, +P_1=1 | — *the reader gives ondole for yondole; no change of count*

2.3 ha-ku-na wo-ko-fu + m-ta-tu-zi u-ya-ta-tu-le (6+9)
lexical = auditory 3·3 + 4·5 — P1=2, P2=5 | P1=3, P2=8 — G 2 + 2 *the reader gives hakuna for hatuna*
2.4 kwa-ko ni ha-fi-fu + ni ma-fu-pi ya-nga-wa ma-le (6+9)
lexical = auditory 2·4 + 4·3?·2 — P1=1, P2=5 | P1=3, P2=6?, P3=8 — G 2 + 3 *the reader gives hafifu, without the Arabic kh*
2.5 mwi-ngi wa lu-tu-fu + mu-wa-wa-zi wa-wi na we-ma (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
Stanza 3 — rhyme a = *ze*, rhyme b = *yo*
3.1 ra-bi tu-a-wa-ze + tu-nga-a-si kwa reh-ma-yo (6+8)
excluded — *guni* of reading: the reader articulates the Arabic syllable *reh*, breaking the syllable count
3.2 si-tu-pa-ti-li-ze + kwa ma-ko-sa tu-ya-ko-sa-yo (6+9)
lexical = auditory 6 + 4·5 — P1=5 | P1=3, P2=8 — G 1 + 2
3.3 tu-pe ba-ra-ka-ze + mu-ha-ma-di ndi-ye tu-mwa-yo (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3 *tumwayo contracts tumwa yako, saving one syllable to fill the nine-syllable slot*
3.4 na wo-te ndu-u-ze + wa-o-mbe-zi wa ma-mbo sa-yo (6+9)
lexical = auditory 3·3 + 4·3·2 — P1=2, P2=5 | P1=3, P2=6, P3=8 — G 2 + 3
3.5 u-tu-nye-she-le-ze + vu-wa nyi-ngi ye-nye re-he-ma (6+9)
lexical 6 + 2·2·2·3 — P1=5 | P1=1, P2=3, P3=5, P4=8 — G-lex 1 + 4
auditory 6 + 4·5 — P1=5 | P1=3, P2=8
— G-aud 1 + 2 — Δ — | −P1=1, −P3=5 *delivered rapidly; only the major stresses are audible*
Stanza 4 — rhyme a = *wa*, rhyme b = *wa*
4.1 tu-le-te-e vu-wa + tu-o-ndo-le-e nyo-ta na ju-wa (6+10)
lexical = auditory 4·2 + 5·2·3 — P1=3, P2=5 | P1=4, P2=6, P3=9 — G 2 + 3 *guni of the text: ten syllables in H2*
4.2 ndi-we mwe-nye ku-wa + u-fi-sha-o na ku-fu-fu-wa (6+9)
lexical = auditory 2·2·2 + 4·5 — P1=1, P2=3, P3=5 | P1=3, P2=8 — G 3 + 2
4.3 ha-li wa-i-ju-wa + hu-hi-ta-ji ku-a-ri-fi-wa (6+9)
lexical = auditory 2·4 + 4·5 — P1=1, P2=5 | P1=3, P2=8 — G 2 + 2
4.4 ha-sha ku-e-me-wa + ni a-mu-ri ku-ni i-ka-wa (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
4.5 wa-la ku-te-nde-wa + ja-la rab-bi mwe-nye a-dha-ma (6+9)
lexical = auditory 2·4 + 2·2·2·3 — P1=1, P2=5 | P1=1, P2=3, P3=5, P4=8 — G 2 + 4 *Arabic realization of rab; no change of count*

Stanza 5 — rhyme a = *si*, rhyme b = *li*
5.1 wa-ngi wa-a-yi-si + kwa ma-te-so ku-wa ta-wi-li (6+9)
lexical 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G-lex 2 + 3
auditory 2·4 + 4·5 — P1=1, P2=5 | P1=3, P2=8 — G-aud 2 + 2 — Δ — | –P2=5
5.2 wa-na wa-si-wa-si + kwa ma-dha-mbi ya-o tha-ki-li (6+9)
lexical = auditory 2·4 + 4·2?·3 — P1=1, P2=5 | P1=3, P2=5?, P3=8 — G 2 + 3
5.3 ra-bi tu-na-fi-si + mu-u-ngwa-na u-so mi-thi-li (6+9)
lexical = auditory 2·4 + 4·2?·3 — P1=1, P2=5 | P1=3, P2=5?, P3=8 — G 2 + 3
5.4 mwi-ngi wa u-kwa-si + tu-pe na-si u-tu-ja-mi-li (6+9)
excluded — false start on *tupe nasi*
5.5 na ya-we me-pe-si + kwa ku-ni ya-ko ka-li-ma (6+8)
lexical = auditory 3·3 + 3·2·3 — P1=2, P2=5 | P1=2, P2=4, P3=7 — G 2 + 3
guni of the text: eight syllables in H2
Stanza 6 — rhyme a = *yo*, rhyme b = *za*
6.1 sa-ya tu-ta-ka-yo + ha-ko m-tu ka-tu mu-we-za (6+9)
lexical = auditory 2·4 + 2·2·2·3 — P1=1, P2=5 | P1=1, P2=3, P3=5, P4=8
— G 2 + 4
6.2 wa-la mwe-nye na-yo + e-la we-we mwe-nye ma-i-za (6+9)
lexical = auditory 2·2·2 + 2·2·2·3 — P1=1, P2=3, P3=5 | P1=1, P2=3, P3=5,
P4=8 — G 3 + 4
6.3 na tu-ta-ma-ni-yo + ku-te-nde-wa ni ye-tu a-za (6+9)
lexical = auditory 6 + 4·3·2 — P1=5 | P1=3, P2=6, P3=8 — G 1 + 3
6.4 kwa yo-te twa-mba-yo + tu-te-nde-e i-li ma-ja-za (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
6.5 tu-shu-ku-ru kwa-yo + i-fu-ra-hi ye-tu mi-ti-ma (6+9)
lexical = auditory 4·2 + 4·2·3 — P1=3, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
Stanza 7 — rhyme a = *di*, rhyme b = *ni*
7.1 mu-o-mba wa-du-di + ha-vi-sha-ka ta-a ya-ki-ni (6+9)
lexical 3·3 + 2·2·2·3 — P1=2, P2=5 | P1=1, P2=3, P3=5, P4=8 — G-lex 2 + 4
auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8
— G-aud 2 + 3 — Δ — | –P1=1
7.2 a-pa-te mu-ra-di + na ma-ta-ko ya-kwe mo-yo-ni (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
7.3 pi-ndi a-ki-na-di + ja a-yu-bu ra-bi ma-sa-ni (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
7.4 ka-tu ha-m-ru-di + m-ja wa-ke mwe-nye ma-a-ni (6+9)
lexical = auditory 2·4 + 2·2·2·3 — P1=1, P2=5 | P1=1, P2=3, P3=5, P4=8
— G 2 + 4 *P1 and P3 minor, P2 and P4 major*

7.5 ha-ta na-m-ru-di + ka-fi-ri-mbi a-so-si-li-ma (6+9)
lexical = auditory 2·4 + 4·5 — P1=1, P2=5 | P1=3, P2=8 — G 2 + 2
Stanza 8 — rhyme a = *mu*, rhyme b = *mu*
8.1 kwa yo-te ma-si-mu + twa-ku-o-mba mo-la ka-ri-mu (6+9)
lexical = auditory 3·3 + 4·2?·3 — P1=2, P2=5 | P1=3, P2=5?, P3=8 — G 2 + 3
8.2 kwa-nza ni a-da-mu + m-tu-me-o na mu-u-nga-mu (6+9)
lexical = auditory 2·4 + 4·5 — P1=1, P2=5 | P1=3, P2=8 — G 2 + 2 *mtumeo contracts mtume wako, to meet the nine-syllable count*
8.3 na wo-te ki-ra-mu + wa-ndi-ka-o ku-la ka-la-mu (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
8.4 lo-ho na ka-la-mu + na zu-o-zo zo-te ka-di-mu (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
8.5 na we-nye ma-si-mu + ma-wa-li-i na wa-ja we-ma (6+9)
lexical = auditory 3·3 + 4·3·2 — P1=2, P2=5 | P1=3, P2=6, P3=8 — G 2 + 3
Stanza 9 — rhyme a = *e*, rhyme b = *ma*
9.1 ra-bi tu-tu-bi-e + ya nu-su-ha ye-nye na-da-ma (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
9.2 u-tu-ghu-fu-ri-e + ku-u dha-mbi i-we la-wa-ma (6+9)
lexical 6 + 2·2·2·3 — P1=5 | P1=1, P2=3, P3=5, P4=8 — G-lex 1 + 4
auditory 6 + 2?·2·5 — P1=5 | P1=1?, P2=3, P3=8
— G-aud 1 + 3 — Δ — | −P3=5 *delivered rapidly*
9.3 na-we u-e-ne-e + u-du-u-ni ya-ko ka-li-ma (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
9.4 tu-ta-wa-si-li-e + kwa i-si-mu ya-ko a-dha-ma (6+9)
lexical = auditory 6 + 4·2?·3 — P1=5 | P1=3, P2=5?, P3=8 — G 1 + 3
9.5 du-a na-i-pa-e + na ka-bu-li i-je kwa hi-ma (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
Stanza 10 — rhyme a = *mu*, rhyme b = *ma*
10.1 du-a na i-ti-mu + kwa ha-mdi na ta-a-dhi-ma (6+8)
lexical = auditory 2·4 + 3·5 — P1=1, P2=5 | P1=2, P2=7 — G 2 + 2 *guni of the text: eight syllables in H2*
10.2 na i-we na-a-mu + na ka-bu-li i-je kwa hi-ma (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
10.3 wa-pa-te na-i-mu + na a-ji-ri we-nye ku-so-ma (6+9)
lexical = auditory 3·3 + 4·2?·3 — P1=2, P2=5 | P1=3, P2=5?, P3=8 — G 2 + 3
10.4 sa-la na sa-la-mu + mba-wa-zi ze-nde-e tu-mwa (6+8)
excluded — *guni* of reading: the reader omits *na*
10.5 na wo-te ki-ra-mu + sa-ha-ba-ze we-nye dhi-ma-ma (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3

Stanza 11 — rhyme a = *li*, rhyme b = *li*
11.1 ra-bi ta-ka-ba-li + na ka-bu-li i-we a-ji-li (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
11.2 yo-te tu-ya-na-li + tu-ta-ka-yo kwa ta-si-hi-li (6+9)
lexical = auditory 2·4 + 4·5 — P1=1, P2=5 | P1=3, P2=8 — G 2 + 2
11.3 ha-tu-na a-ma-li + tu-te-nda-ye ni-ka-bu-li (6+8)
excluded — *guni* of reading: false start on *tutenda*, and the text is distorted
11.4 e-la kwa a-ji-li + na re-he-ma za-ko ja-la-li (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
11.5 ndi-yo ta-wa-su-li + ma-o-mbe-zi ye-tu wa-fa-ma (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
Stanza 12 — rhyme a = *ni*, rhyme b = *ri*
12.1 twa-ta-ka a-ma-ni + ya gha-dha-bu za-ko ja-ba-ri (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
12.2 hu-ku du-ni-ya-ni + na ki-ya-ma si-ku ya na-ri (6+9)
lexical = auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
12.3 tu-ti-e pe-po-ni + tu-sa-ki-ni ka-ti-ka da-ri (6+9)
lexical = auditory 3·3 + 4·3·2 — P1=2, P2=5 | P1=3, P2=6, P3=8 — G 2 + 3
12.4 u-tu-pe hi-sa-ni + wa-ke we-ma na ku-khi-ta-ri (6+9)
lexical = auditory 3·3 + 2·2·5 — P1=2, P2=5 | P1=1, P2=3, P3=8 — G 2 + 3
12.5 na-o wil-da-ni + wa-tu-pa-mbe mbe-le na nyu-ma (5+9)
lexical = auditory 2·3 + 4·2·3 — P1=1, P2=4 | P1=3, P2=5, P3=8 — G 2 + 3
guni of the text: five syllables in H1, the reader articulating wil
Stanza 13 — rhyme a = *mba*, rhyme b = *za*
13.1 na mwe-nye ku-o-mba + ku-te-nde-wa ni ya-kwe ja-za (6+9)
lexical = auditory 3·3 + 4·3·2 — P1=2, P2=5 | P1=3, P2=6, P3=8 — G 2 + 3
13.2 ha-ta mwe-nye kwi-mba + ja-za ya-kwe ya-wa ku-tu-zwa (6+9)
excluded — false start on *yawa*
13.3 wa-lla-hi 'mi na-mba + a-o-mba-ye mo-la a-zi-za (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
13.4 ku-la ma-ti-la-ba + yu-wa-pe-wa na nyi-ngi ne-ma (6+9)
lexical = auditory 2·4 + 4·3·2 — P1=1, P2=5 | P1=3, P2=6, P3=8 — G 2 + 3
13.5 na tu-ta-ma-ni-yo + ku-te-nde-wa ni ye-tu a-za (6+9)
lexical 6 + 4·3·2 — P1=5 | P1=3, P2=6, P3=8 — G-lex 1 + 3
auditory 6 + 4·5 — P1=5 | P1=3, P2=8 — G-aud 1 + 2 — Δ — | −P2=6
Stanza 14 — rhyme a = *nga*, rhyme b = *ni*
14.1 na-mi ni-nga-tu-nga + si-ya-ju-i ni ku-ta-ma-ni (6+9)
lexical = auditory 2·4 + 4·5 — P1=1, P2=5 | P1=3, P2=8 — G 2 + 2

14.2 si-ju-i ku-pa-nga + lu-lu nge-ma na ma-ri-ja-ni (6+9)
lexical = auditory 3·3 + 2·2·5 — P1=2, P2=5 | P1=1, P2=3, P3=8 — G 2 + 3
14.3 ha-ta-u-wa sha-nga + ma-ku-ru-to ya-so tha-ma-ni (6+9)
lexical 4·2 + 4·2·3 — P1=3, P2=5 | P1=3, P2=5, P3=8 — G-lex 2 + 3
auditory 2·4 + 4·2·3 — P1=1, P2=5 | P1=3, P2=5, P3=8
— G-aud 2 + 3 — Δ −P1=3, +P1=1 | —
14.4 kwa nyo-te ma-le-nga + mu-o-na-po pe-nye la-ha-ni (6+9)
lexical = auditory 3·3 + 4·2·3 — P1=2, P2=5 | P1=3, P2=5, P3=8 — G 2 + 3
14.5 ha-mu-wi wa-ji-nga + mu-fa-nya-po mwa-pa-ta me-ma (6+9)
excluded — false start on *mufanya*

2.b. Excluded lines

Six lines of the seventy mapped are removed from the statistical pool. Four are *guni* of reading, in which the reader adds, omits or distorts a syllable the source text does not contain, and two are false starts, which cannot be verified.

3.1 *ra-bi tu-a-wa-ze* + *tu-nga-a-si kwa reh-ma-yo* — *guni* of reading: the reader articulates the Arabic syllable *reh*, breaking the syllable count
5.4 mwi-ngi wa u-kwa-si + tu-pe na-si u-tu-ja-mi-li — false start on tupe nasi
10.4 sa-la na sa-la-mu + mba-wa-zi ze-nde-e tu-mwa — *guni* of reading: the reader omits na
11.3 *ha-tu-na a-ma-li* + *tu-te-nda-ye ni-ka-bu-li* — *guni* of reading: false start on *tutenda*, and the text is distorted
13.2 ha-ta mwe-nye kwi-mba + ja-za ya-kwe ya-wa ku-tu-zwa — false start on yawa
14.5 ha-mu-wi wa-ji-nga + mu-fa-nya-po mwa-pa-ta me-ma — false start on mufanya

2.c. Divergences between the source text and the execution

Nine lines of the seventy diverge from the text supplied to the reader. Four are removed from the pool as *guni* of reading and are listed above. The five retained are metrically inert: at 2.2 the reader gives *ondole* for *yondole*, at 2.3 *hakuna* for *hatuna*, at 2.4 *hafifu* for *khafifu* without the Arabic articulation, at 4.5 *rabbi* with the Arabic geminate, and at 7.1 *havishaka taa* for *havi shaka twaa*. None alters the syllable count, and the lexical reading is derived in every case from the source and not from the execution.

2.d. Summary

The pool comprises 64 lines of the 70 mapped, that is 128 hemistiches.

Condition 1. The matrix requires 6 + 9 in every line. The constants hold in **57 lines of 64, that is 89.1 percent**. The seven exceptions are *guni* of the source text: H2 carries seven syllables at 1.1, eight at 1.4, 5.5 and 10.1, and ten at 2.1 and 4.1, while H1 carries five at 12.5, the reader articulating *wil* in *wildani*. All seven are retained in the pool and count against the matrix.

The lengths present in this corpus. Of the 128 retained hemistiches, 63 measure six syllables and 58 measure nine; the seven remaining measure five, seven, eight or ten. **Three hemistiches of 128, that is 2.3 percent, measure eight syllables**, and all three are *guni* of the text. Measured in syllables, 24 of the 956 in the corpus, that is 2.5 percent, lie in a segment whose length is a multiple of four.

Distribution of G-lex, derived from the printed text. In the six-syllable hemistich, G = 2 in 56 lines, G = 1 in 6 and G = 3 in 2. In the nine-syllable hemistich, G = 3 in 46 lines, G = 2 in 9 and G = 4 in 9.

Distribution of G-aud, established by ear. In the six-syllable hemistich, G = 2 in 56 lines, G = 1 in 6 and G = 3 in 2. In the nine-syllable hemistich, G = 3 in 44 lines, G = 2 in 14 and G = 4 in 6.

Transition. The reading departs from the lexical structure in **10 hemistiches of 128, that is 7.8 percent**. The departures are declared individually in the mapping above.

The sequence 4·4 occurs once in the whole corpus on the auditory measure, at 1.4, and that line is one of the *guni* of the text: its second hemistich carries eight syllables where the matrix prescribes nine. On the lexical measure it does not occur at all.

Periodicity. The nuclei of a hemistich are all of equal volume in 24 hemistiches of 128 on the lexical measure, that is 18.8 percent, and in 23 on the auditory measure, that is 18.0 percent.

Inter-pass control. No inter-pass comparison is reported for this corpus. The first annotation was carried out under a notation and a set of conventions that preceded those of the present study, and it did not segment the second hemistich in the terms used here; a coincidence rate between the two would measure the change of instrument rather than the stability of the listener. The comparison is made on the two *shairi* corpora that follow, where both annotations worked under the same conventions.

3. Statistical Analysis and Falsification

4.a. What is tested here, and what is not

This corpus is not equivalent to the *utenzi* corpora, and the difference is not one of degree. It must be stated before any figure is given, because it determines which conditions can be applied at all.

The line of this poem measures fifteen syllables, divided 6 + 9. **Neither hemistich is a container of eight syllables**, and of the 128 retained hemistiches only three measure eight, all three being *guni* of the source text.

This matters because the accentual model is not formulated for a container of any length. Vierke's bipartite axiom states that *the eight-syllable line* carries two accentual groups; her *Grenzschema* specifies $|A_1| = 4$, $|A_2| = 4$ with stresses on positions 3 and 7, which presupposes a segment of eight. Minerba extends the derivation to the *shairi* on the ground that the meter adopts "syllabic lengths of multiples of four" (Minerba, 2025, p. 217), and he analyses the *shairi* line as a reduplication of the eight-syllable *kipande* of the *utenzi* (p. 219).

The *Waji Waji* is a classical nineteenth-century *shairi*, and it offers no such segment. Conditions 5, 6 and 8 are therefore applied to it, but the outcome is of a different kind from that of the preceding case studies, and the section says so where it arises.

Conditions 1, 2 and 3 are applied to the line and to the stanza, as elsewhere. **Condition 4 is applicable**, this corpus employing the permutation *kupindua kina kimoja*.

4.b. Conditions 1, 2 and 4

Condition 1. The constants 6 + 9 hold in **57 lines of 64, that is 89.1 percent**. The seven exceptions are *guni* of the source text and are of both kinds — six of excess or deficiency in H2, one of deficiency in H1 — declared line by line in section 3, retained in the pool, and counted against the matrix. No exception arises from the execution: every line whose count departs from the source was removed before counting.

Condition 2. The local rhyme *a* binds the first hemistich of the first four lines of each stanza; the rhyme *b* binds their second hemistich. Both hold in their syllabic form, a^s, across the fourteen stanzas, and the vocalic form is not invoked. The global constant

closes the fifth line of every stanza on *-ma*, and it holds in **14 stanzas of 14**. X^{s} is satisfied and X^{v} is not required.

Condition 4. The matrix specifies the permutation *kupindua kina kimoja*: in the fifth line of each stanza, the rhyme *b* that closed the preceding lines moves to the caesura, and the terminal position passes to the global constant X. The permutation is executed in **14 stanzas of 14**, without a single failure of the chain. The composer relocates a rhyme he has established over four lines, and does so fourteen times.

Condition 3. No word crosses the caesura in any of the 64 retained lines.

Verdict. The syllabic matrix describes this corpus with seven declared *guni*, and the rhyme architecture — two local chains, one permutation and one global constant — holds without exception. The architecture is the more demanding of the two tested in this chapter, and it is the one that holds.

4.c. Condition 5: the bipartite axiom

The axiom concerns the eight-syllable segment. This corpus contains three, and they are defective lines. The condition is therefore applied to the segments the corpus does supply, and the result is reported by hemistich rather than consolidated, because the consolidation conceals what matters.

In the **six-syllable hemistich**, the lexicon yields G = 2 in **56 lines of 64**, G = 1 in six and G = 3 in two, and the reading realizes the same distribution exactly.

In the **nine-syllable hemistich**, the lexicon yields G = 3 in **46 lines of 64**, G = 2 in nine and G = 4 in nine; the reading yields G = 3 in 44, G = 2 in fourteen and G = 4 in six.

Consolidated, the figures give G = 2 in 65 hemistiches of 128 on the lexical measure and 70 on the auditory one, that is 50.8 and 54.7 percent — proportions indistinguishable from those reported for the *utenzi* and the *wimbo*. Separated, they show something the consolidated figure hides.

The number of accentual groups follows the length of the container. A segment of six syllables accommodates two nuclei and rarely more, because the Swahili lexicon inside six syllables rarely yields three word-groups; a segment of nine accommodates three, for the same reason in reverse. The same speaker, in the same line, produces two groups in the first half and three in the second, one hundred and twenty-eight times over.

The observation is not new, and the first scholar to count accents in Swahili verse had already made it. Sacleux does not assign a fixed number of accents to the Swahili line. He assigns a number to each verse length: two for lines of three, four, six and eight syllables, three for the ten-syllable line, four for those of twelve and fourteen (1909, p. 327). The nine-syllable line does not appear in his list, but his own progression brackets it, and the corpus falls where the bracket predicts. His figure for the six-syllable line is verified here with precision: two accentual groups in 56 of the 64 first hemistiches, that is 87.5 percent.

What the accentual school retained from Sacleux was the count for one length. What it set aside was the fact that he made the count depend on the length at all — which is to say, the fact that the number of accents is a property of the container and not of the meter.

Verdict. The axiom is not falsified on this corpus so much as rendered inapplicable to it. Where it can be tested at all — on the three defective lines that happen to carry eight syllables — it holds in one and fails in two. What the corpus shows is more general: G is a function of segment length and of Swahili word structure, and a model that fixes it at two for one particular length has described an arithmetical consequence and called it a law.

4.d. Condition 6: the *Grenzschema*

The ideal configuration requires two tetrasyllabic nuclei with stresses on positions 3 and 7, in a segment of eight syllables.

The sequence 4·4 does not occur once in the entire corpus on the lexical measure, and occurs once on the auditory one, at 1.4 — in a line whose second hemistich carries eight syllables where the matrix prescribes nine, and where the reading fuses into two tetrasyllabic nuclei what the lexicon gives as 2·2·2·2. The single occurrence of the ideal configuration in this classical *shairi* is thus produced by a defect of the text and by a fusion in the reading, and by nothing in the composition.

Two further occurrences are worth recording, and they arise from the reading rather than from the text.

At 3.1 the reader articulates the Arabic syllable *reh* in *rehemayo*, which breaks the syllable count and yields a second hemistich of eight syllables partitioned 4·4 on positions 3 and 7. At 11.3 a false start followed by a distortion of the text produces the same outcome: a hemistich of eight syllables, 4·4, positions 3 and 7. Both lines are excluded from the pool as *guni* of reading.

The observation deserves to be stated plainly, because it bears on what the *Grenzschema* is. In this corpus the ideal configuration of the accentual school appears three times: once in a defective line of the source, and twice in lines the reader deformed. It never appears in a regular line, for the arithmetical reason that a regular line has no eight-syllable segment. What the three occurrences share is not a metrical intention but an accident that produced a segment of eight syllables in which two words of four syllables happened to fall together. The configuration is a phonetic consequence of that arithmetic, and the three cases show it in isolation, unaccompanied by anything metrical whatever.

Verdict. The prototype does not govern this corpus and cannot: the segment it specifies does not occur in a regular line of the poem.

4.e. Condition 7: periodicity

The nuclei of a hemistich are all of equal volume in **24 hemistiches of 128 on the lexical measure, that is 18.8 percent**, and in 23 on the auditory one. In the remaining hemistiches no unit repeats.

On the lexical measure the six-syllable hemistich attests 2·4 in 29 lines, 3·3 in 21, a single nucleus of six in six, 4·2 in five, 2·2·2 in two and 2·3 in one; the nine-syllable hemistich attests 4·2·3 in 30 lines, 4·3·2 in nine, 2·2·2·3 in eight, 4·5 in eight, 2·2·5 in three and six further sequences once each. The first stressed syllable falls on positions 1 to 3 in the six-syllable hemistich and on positions 1 to 4 in the nine-syllable one.

Verdict. There is no repeating unit in four hemistiches of five, and the two hemistiches of a single line never carry the same sequence, since their lengths differ. No periodicity is available at either scale.

4.f. Condition 8: the tetrasyllabic foot

This is where the corpus is decisive, and the decision requires no measurement.

Neither six nor nine is a multiple of four. A segment of six admits 4 + 2 or 2 + 4 but not two tetrasyllabic units; a segment of nine admits 4 + 5 or 4 + 2 + 3 but never a partition into four-syllable feet. Of the 128 retained hemistiches, **three have a length that is a multiple of four, that is 2.3 percent**, and all three are *guni* of the text. Measured in syllables, **24 of 956, that is 2.5 percent**, lie in a segment the imported template could describe.

A tetrasyllabic nucleus of some kind occurs in 82 hemistiches on the lexical measure, that is 64.1 percent, and the frequency is unsurprising: Swahili produces four-syllable words in abundance. But **in no hemistich of this corpus do tetrasyllabic nuclei exhaust the segment**, because the arithmetic forbids it. The template appears as one unit among others, flanked by units of two, three or five, and it never tiles a line.

Verdict. For 97.7 percent of the metrical units of this poem, the derivation from an imported tetrasyllabic template produces no prediction at all. This is not a statistical failure. The model has no point of application on a classical *shairi* of the nineteenth century, while being asserted of the *shairi* as a meter.

4.g. What the corpus establishes

Two models have been applied to the same 64 lines.

The syllabic matrix predicts the syllable count in 89.1 percent of the lines, accounts for the remaining 10.9 percent through the *guni* the tradition itself defines, and predicts the rhyme architecture — two local chains, a permutation executed fourteen times, and a global constant — without a single exception. The reading departs from the lexical accentual structure in 10 hemistiches of 128, that is 7.8 percent, against 12.5 percent in the prose control corpus of section 9.7.

The accentual matrix predicts nothing here. Its axiom is stated for a segment of eight syllables that this poem does not contain; its prototype occurs three times, always in a line that is defective in the text or in the reading; its template is arithmetically excluded from 97.7 percent of the segments.

The point is not that the accentual model performs badly on this corpus. It is that a classical *shairi*, composed in the nineteenth century by a poet of the Swahili language, in a form the tradition names and Shariff documents, falls entirely outside the reach of a model asserted to govern the *shairi* meter. A theory that cannot describe a canonical instance of the form it claims to explain has not been contradicted by the data. It has been shown to have no data.

What the corpus shows in its place is a construct of another order. A fifteen-syllable line divided 6 + 9, a five-line stanza, two local rhyme chains, a permutation of the terminal rhyme executed in every stanza, and a global constant holding across fourteen stanzas: the architecture is arithmetically locked at every point, and the composer meets it in 89.1 percent of his lines, accounting for the remainder through a category his own tradition names. The system this corpus exhibits is

not a simplified or a deviant form of anything. It is a sui generis construct, governed by syllable and rhyme, and it is what the composer of the nineteenth century worked under — not the bed to which an exogenous theory would fit him.

9.6.2. Case Study VII: Test of Continuity of the Classical 16-Syllable "Kupindua Kina Kimoja" Matrix — *Kiswahili!!*, by Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau) (2003)

The model tested here is the one formalized in section 2.5.2. Minerba derives the classical Swahili line from an imported tetrasyllabic template and states the derivation for the *utenzi*, the *shairi* and the *wimbo* alike (Minerba, 2025, p. 217). The model operates at two degrees of strictness, and both are tested: in its strict form it requires two tetrasyllabic nuclei with stresses on positions 3 and 7, which is the *Grenzschema*; in its extended form it requires only that the first stress fall on an odd syllable (Minerba, 2025, p. 220).

Sacleux had already given a different account. He assigns a number of accents to each verse length rather than a fixed count for the meter: two accents for lines of three, four, six and eight syllables, three for the ten-syllable line (1909, p. 327), and he defines the accent that carries this count as free, in both placement and character, from the tonic stress proper to each word (1909, p. 8). Minerba inherits Vierke's framework and does not retain this dependency on length.

The corpus is the modern classical *shairi Kiswahili!!*, composed in Kiamu by Sheikh Mahmud Ahmed Abdull Kadir, Ustadh Mau, in 2003. For this corpus the procedure runs in the opposite direction to the others. The poet read his composition aloud in Lamu in September–October 2005 (Roy, 2013a, p. 10), from a printed sheet in Roman script headed *Kiswahili !!*; the reading was recorded, transcribed in IPA (Roy & ibn Kadir, 2013), and mapped, and the text studied here is the one that results. No text was supplied to a reader and then compared with its execution: the reader is the poet, and his reading is the text. The recording is deposited in open access (Roy & ibn Kadir, 2016), so that the analysis may be reproduced independently. The stanza numbering follows the version the poet published in 2013 (Abdulkadir & Frankl, 2013), which differs from the text read in 2005 in the places listed at the end of section 3.

The recording captures the poem formally read by its own author, and not sung. This sets aside Layer 3, as section 9.3 establishes: what

the recording gives access to is not a layer of the metrical system, since Layers 1 and 2 say nothing about accent. Accent belongs to the language, and the reading allows the extent to which the lexical structure is realized to be observed, nothing more.

The syllabification and the accentual mapping follow the three-step protocol of section 9.2.3, restated in the preliminaries to this section. The accentual structure of each line is first derived from the printed text by the penultimate rule; the formal reading is then screened by ear, and the two results are reported side by side. Acoustic extraction is applied to a sample of hemistiches only, and it is reported for the three *shairi* corpora together in section 9.6.4.

1. Identification of the Formal Syllabic Matrix

The generic matrix available at this stage of the argument, QRG.4.1, is the relaxed bipartite form, the 16-syllable *shairi* meter, specifically deploying the generative sub-meter known as ***Kupindua kina kimoja*** (the permutation of a single rhyme).

Formal Parameters:

$n_{max} = 4, H = 2, S = 16$(where $S = 8 + 8$).

The Matrix:

L_1: $[8, a] + [8, b]//$
L_2: $[8, a] + [8, b]//$
L_3: $[8, a] + [8, b]//$
L_4: $[8, b] + [8, X]//$

The Algebraic Equation for this specific stanzaic configuration:

$$\boldsymbol{Shairi}_{\boldsymbol{Kupindua_kina_kimoja}} = \mathbf{3}([\mathbf{8}\boldsymbol{a}] + [\mathbf{8}\boldsymbol{b}]//) \cdot ([\mathbf{8}\boldsymbol{b}] + [\mathbf{8}\boldsymbol{X}]//)$$

Across the stanzas, the first three lines share an internal rhyme *a* and a terminal rhyme *b*, which act as local variables changing from one stanza to the next. In the fourth line, the preceding terminal rhyme *b* permutes to the position of the internal caesura, while the absolute terminal boundary is locked by the global constant rhyme *X* extending across the entire poem.

This form is satisfied by *Kiswahili!!* without exception across all eighteen stanzas: the local rhyme *a* of H1 changes every stanza and is shared by L1–L3; the rhyme *b* closing H2 of L1–L3 recurs at the caesura of L4; a further rhyme *X* closes L4.

Nonetheless, the corpus imposes a stricter relation than the generic matrix requires, and it should be stated rather than left implicit. In *Kiswahili!!*, the rhyme *b* that permutes to the caesura of L4 is not a variable renewed each stanza, as it is in the *Waji Waji*: it is the **same syllable, -ni, across all eighteen stanzas**. What the relaxed matrix treats as a stanza-local variable is here a second global constant, locked to *both* the terminal position of H2 in L1–L3 and the caesura position of L4. The corpus therefore realizes the stricter form:

The Matrix:

L_1: $[8, a] + [8, B]//$
L_2: $[8, a] + [8, B]//$
L_3: $[8, a] + [8, B]//$
L_4: $[8, B] + [8, X]//$

The Algebraic Equation:

$$\boldsymbol{Shairi}_{\boldsymbol{Kupindua_kina_kimoja,global}} = \mathbf{3}([\mathbf{8}\boldsymbol{B}] + [\mathbf{8}\boldsymbol{B}]//) \cdot ([\mathbf{8}\boldsymbol{B}] + [\mathbf{8}\boldsymbol{X}]//)$$

Two global constants, not one, govern this poem: *B* (*-ni*) and *X* (*-ta*), each in its syllabic form. This is a more demanding architecture than the generic matrix anticipates, and it is realized by a poet whose Swahili is native.

2. Exhaustive Syllabic and Accentual Mapping

The syllabification and the accentual structure are mapped across the whole corpus, hemistich by hemistich, in the notation set out in the preliminaries to this section: the sequence of nucleus volumes, the positional index Pi of each nucleus, and G, the number of accentual groups. Standard Swahili orthography is retained throughout.

The auditory reading for this corpus was cross-checked against an earlier IPA transcription made independently, by ear alone and prior to any acoustic analysis, in collaboration with the linguist Marie-Françoise Rombi. Where the two disagreed, the recording was reheard and the auditory reading given below is the one retained.

3. Kiswahili!!: *Syllabic and Accentual Mapping, Stanzas 1–18*

Each line is given as two hemistiches separated by +, in the notation of the preliminaries to this section. Where the lexical and auditory readings coincide, a single line gives both; where they differ, the two are given separately, followed by Δ.

Stanza 1

1.1 ku-nya-ma-a ni-me-cho-ka + tʰa-nya-ma-a ha-ta li-ni (8+8)
lexical 4·4 + 4·2·2 — P1=3, P2=7 | P1=3, P2=5, P3=7 — G-lex 2 + 3
auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G-aud 2 + 2 — Δ — | –P2=5
1.2 wa-na-ngu hu-ni-e-pu-ka + ku-wa-o-na na-ta-ma-ni (8+8)
lexical = auditory 3·5 + 4·4 — P1=2, P2=7 | P1=3, P2=7 — G 2 + 2
1.3 wa-lo-ba-ki ku-ni-shi-ka + si-wa-ngu ni-wa-we-nda-ni (8+8)
lexical = auditory 4·4 + 3·5 — P1=3, P2=7 | P1=2, P2=7 — G 2 + 2
1.4 mi-mi ni-me-wa-te-nda-ni + mbo-na mwa-ni-pi-ja zi-tʰa (8+8)
H1 excluded — H1 false start on *nimewatendani*; H2's middle nucleus (would fall at position 5) judged too weak to count, P2=7 clearly major
lexical H2 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory H2 2·6 — P1=1, P2=7 — G-aud 2 — Δ –P2=5

Stanza 2

2.1 wa-na-ngu mi-mi wa-da-mu + wa-na wa u-swa-hi-li-ni (8+8)
lexical = auditory 3·2·3 + 2·6 — P1=2, P2=4, P3=7 | P1=1, P2=7 — G 3 + 2
2.2 a-si-li ha-wa-na ha-mu + ya-ku-ni-yu-wa ni-na-ni (8+8)
lexical = auditory 3·3?·2 + 5·3 — P1=2, P2=5?, P3=7 | P1=4, P2=7 — G 3 + 2
2.3 wa-me-na-ti-ya ka-u-mu + na wa-na wa ma-ji-ra-ni (8+8)
lexical = auditory 5·3 + 3·5 — P1=4, P2=7 | P1=2, P2=7 — G 2 + 2
2.4 ko-sa la-ngu ko-sa ga-ni + mbo-na hu-ni-pi-ja zi-tʰa (8+8)
lexical = auditory 2?·2·2?·2 + 2·4·2 — P1=1?, P2=3, P3=5?, P4=7 | P1=1, P2=5, P3=7 — G 4 + 3

Stanza 3

3.1 mi-mi ma-me-nu si-tʰa-sa + wa-la si-na pu-ngu-wa-ni (8+8)
lexical = auditory 2·3·3 + 2?·2·4 — P1=1, P2=4, P3=7 | P1=1?, P2=3, P3=7 — G 3 + 3 *reader articulates aspirated tʰa, unnoticed in the first pass*
3.2 ni-me-za-a wa-ma-mba-sa + na ku-ngi-ne zi-si-wa-ni (8+8)
lexical = auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G 2 + 2

3.3 ni-ze-e wa-na-si-ya-sa + na zi-yo-ngo-zi wa-di-ni (8+8)
lexical 3·5 + 5·3 — P1=2, P2=7 | P1=4, P2=7 — G-lex 2 + 2
auditory 3·2?·3 + 5·3 — P1=2, P2=4?, P3=7 | P1=4, P2=7
— G-aud 3 + 2 — Δ +P2=4 | — *wanasiasa is a compound in which each component carries its own stress*
3.4 ma-fu-ndi wa ku-la fa-ni + na ma-shu-ja-a wa-zi-ta (8+8)
lexical = auditory 3·3?·2 + 5·3 — P1=2, P2=5?, P3=7 | P1=4, P2=7 — G 3 + 2
Stanza 4
4.1 ndi-mi ma-ma-ke mu-ya-ka + pi-ya mwe-ngo a-thu-ma-ni (8+8)
lexical = auditory 2?·3·3 + 2?·2·4 — P1=1?, P2=4, P3=7 | P1=1?, P2=3, P3=7
— G 3 + 3
4.2 na za-hi-di ka-dha-li-ka + na we-ngi wa-ke we-nda-ni (8+8)
lexical = auditory 4·4 + 3·2·3 — P1=3, P2=7 | P1=2, P2=4, P3=7 — G 2 + 3
4.3 a-li kʰo-ti na ma-ta-ka + wo-te mbwa mo-ya ka-ri-ni (8+8)
lexical 2·2·4 + 2·3·3 — P1=1, P2=3, P3=7 | P1=1, P2=4, P3=7 — G-lex 3 + 3
auditory 4·4 + 2·3·3 — P1=3, P2=7 | P1=1, P2=4, P3=7
— G-aud 2 + 3 — Δ −P1=1 | —
4.4 wa-li-to-ka ma-tʰu-mbo-ni + wa ka-wa-a ka-ma nyo-tʰa (8+8)
lexical 4·4 + 4·2·2 — P1=3, P2=7 | P1=3, P2=5, P3=7 — G-lex 2 + 3
auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G-aud 2 + 2 — Δ — | −P2=5
Stanza 5
5.1 i-nki-sha-fi nga-li-ya + u-ki-so-me na ki-da-ni (7+8)
lexical = auditory 4·3 + 4·4 — P1=3, P2=6 | P1=3, P2=7 — G 2 + 2 *guni of the text: seven syllables in H1, the reader not articulating n and ga separately*
5.2 ndi-po ta-ka-po kwe-le-ya + ni-kwa-mbi-ya-o mwe-nda-ni (8+8)
lexical = auditory 2?·3?·3 + 5·3 — P1=1?, P2=4?, P3=7 | P1=4, P2=7
— G 3 + 2
5.3 ni tʰu-ngo zi-me-ba-ki-ya + na ha-zi-fi a-si-la-ni (8+8)
lexical = auditory 3·5 + 4·4 — P1=2, P2=7 | P1=3, P2=7 — G 2 + 2 *the poet reads* zimebakiya*, "the compositions have remained"; the 2013 edition prints* zimesaliya
5.4 wa-lo-zi-tʰu-nga ni nya-ni + ni wa-na-ngu wa-lo-pi-ta (8+8)
lexical = auditory 5?·3 + 4?·4? — P1=4?, P2=7 | P1=3?, P2=7?
— G 2 + 2 *H2 stresses minor relative to H1*
Stanza 6
6.1 na ma-le-nga wa m-vi-ta + na pi-ya chi-ra-ghu-di-ni (8+8)
lexical = auditory 4·4 + 3·5 — P1=3, P2=7 | P1=2, P2=7 — G 2 + 2

6.2 nya-yo wa-li-zi-fu-wa-ta + ha-wa-ku-ki-ri u-du-ni (8+8)
lexical = auditory 2·6 + 5·3 — P1=1, P2=7 | P1=4, P2=7 — G 2 + 2
6.3 n-na-ba-ha-ni hu-tʰe-tʰa + la-ki-ni hu-fa-li-ya-ni (8+8)
lexical = auditory 5·3 + 3·5 — P1=4, P2=7 | P1=2, P2=7 — G 2 + 2
6.4 ndi-ye pwe-ke u-wa-nda-ni + i-nga-wa a-me-i-kʰi-ta (8+8)
lexical = auditory 2·2·4 + 3·5 — P1=1, P2=3, P3=7 | P1=2, P2=7 — G 3 + 2

Stanza 7

7.1 ba-do ku-za-a na-we-za + si-ya-ko-ma u-ki-ngo-ni (8+8)
lexical = auditory 2·3·3 + 4·4 — P1=1, P2=4, P3=7 | P1=3, P2=7 — G 3 + 2
7.2 la-ki-ni mu-me-ni-pu-za + mu-me-i-to-wa fu-wo-ni (8+8)
lexical = auditory 3·5 + 5·3 — P1=2, P2=7 | P1=4, P2=7 — G 2 + 2 *the poet reads* puza*; the 2013 edition spells* puuza*, which would give nine syllables*
7.3 wa-ngi-ne me-i-to-ke-za + ku-ni-pa-ngi-ya ka-nu-ni (8+8)
lexical = auditory 3·5 + 5?·3 — P1=2, P2=7 | P1=4?, P2=7 — G 2 + 2
7.4 mu-sa-mi-ya-ti ku-bu-ni + nyi-nyi mu-li-po-ni-wa-ta (8+8)
lexical = auditory 5·3 + 2·6 — P1=4, P2=7 | P1=1, P2=7 — G 2 + 2

Stanza 8

8.1 hu-li-ya ki-si-ki-ti-ka + cha-nga-li-ya ja-ri-da-ni (8+8)
lexical = auditory 3·5 + 4·4 — P1=2, P2=7 | P1=3, P2=7 — G 2 + 2
8.2 we-ngi wa-na-o-a-ndi-ka + si wa-na-ngu ni wa-ge-ni (8+8)
lexical = auditory 2·6 + 4·4 — P1=1, P2=7 | P1=3, P2=7 — G 2 + 2
8.3 i-dha-a-ni ka-dha-li-ka + wa-pe-ka tʰu-ngo ni nya-ni (8+8)
lexical = auditory 4·4 + 3?·2·3 — P1=3, P2=7 | P1=2?, P2=4, P3=7 — G 2 + 3
8.4 we-ngi ha-wa-to-ki pwa-ni + li-cha ku-wa mbwa m-vi-ta (8+8)
lexical = auditory 2·4?·2 + 2·2·4 — P1=1, P2=5?, P3=7 | P1=1, P2=3, P3=7 — G 3 + 3

Stanza 9

9.1 a-nga-li-ya na zi-ta-bu + zi-so-me-shwa-o shu-le-ni (8+8)
lexical = auditory 4·4 + 5·3 — P1=3, P2=7 | P1=4, P2=7 — G 2 + 2
9.2 ha-za-ndi-kwi na ra-ja-bu + si su-di wa-la si sha-ni (8+8)
lexical = auditory 4·4 + 3·2?·3 — P1=3, P2=7 | P1=2, P2=4?, P3=7 — G 2 + 3
9.3 njo-ro-ge ndi-yo ka-ti-bu + a-shi-shi-ye-o su-ka-ni (8+8)
lexical = auditory 3·2?·3? + 5·3 — P1=2, P2=4?, P3=7? | P1=4, P2=7 — G 3 + 2
9.4 cha-ro na wa-ke we-nda-ni + na-o nyu-ma hu-fu-wa-ta (8+8)
lexical = auditory 2·3·3 + 2?·2·4 — P1=1, P2=4, P3=7 | P1=1?, P2=3, P3=7 — G 3 + 3

Stanza 10

10.1 hu-a-li-kwa ko-nga-ma-no + che-nda hu-ru-di ndi-ya-ni (8+8)
lexical = auditory 4·4 + 2·3·3 — P1=3, P2=7 | P1=1, P2=4, P3=7 — G 2 + 3

10.2 hu-o-na u-tu-ngu m-no + ku-wa nyi-nyi si-wa-o-ni (8+8)
lexical = auditory 3?·3?·2 + 2?·2?·4 — P1=2?, P2=5?, P3=7 | P1=1?, P2=3?, P3=7 — G 3 + 3

10.3 na hu-zi-u-ma zi-ta-no + la-ki-ni ni-tʰe-nde ni-ni (8+8)
lexical = auditory 5·3 + 3·3·2 — P1=4, P2=7 | P1=2, P2=5, P3=7 — G 2 + 3

10.4 wa-na-ngu mu-me-i-khi-ni + ma-me-nu mu-me-ni-wa-ta (8+8)
H1 excluded — H1 false start on *khini*
lexical = auditory H2 3·5 — P1=2, P2=7 — G 2

Stanza 11

11.1 na hu-li-ya kwa ma-to-zi + cha-nga-li-ya mi-ti-ha-ni (8+8)
lexical = auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G 2 + 2

11.2 wa-na-fu-ndi wa ki-bwe-zi + na wa ki-su-mu zi-wa-ni (8+8)
lexical = auditory 4·4 + 5·3 — P1=3, P2=7 | P1=4, P2=7 — G 2 + 2

11.3 ndi-wo wa-na-o ba-ri-zi + wa-li-yo-ko ki-le-le-ni (8+8)
lexical 2·3·3 + 4·4 — P1=1, P2=4, P3=7 | P1=3, P2=7 — G-lex 3 + 2
auditory 2·6 + 4?·4 — P1=1, P2=7 | P1=3?, P2=7
— G-aud 2 + 2 — Δ −P2=4 | —

11.4 mu-lo-to-ka kwe-tʰu pwa-ni + mu-ko tʰi-ni hu-ko-ko-ta (8+8)
lexical = auditory 4·2?·2 + 2?·2·4? — P1=3, P2=5?, P3=7 | P1=1?, P2=3, P3=7? — G 3 + 3

Stanza 12

12.1 wa-fa-nya-o u-ta-fi-ti + wa di-gri-i zu-wo-ni (8+7)
lexical = auditory 4·4 + 4·3 — P1=3, P2=7 | P1=3, P2=6 — G 2 + 2 *guni of the text as read: seven syllables in H2; the 2013 edition prints wa uzamili, eight syllables*

12.2 wa-swa-hi-li ni ka-ti-ti + a-u ha-wa-pa-ti-ka-ni (8+8)
lexical = auditory 4·4 + 2·6 — P1=3, P2=7 | P1=1, P2=7 — G 2 + 2

12.3 ni nya-ni ni-m-la-i-ti + mwe-nye ma-ko-sa ni nya-ni (8+8)
lexical = auditory 3·5 + 2?·3?·3 — P1=2, P2=7 | P1=1?, P2=4?, P3=7 — G 2 + 3

12.4 mi-mi ha-mu-ni tha-mi-ni + m-ngi-ne ha-mu-ku-pa-tʰa (8+8)
H2 excluded — H2 false start on *mngine*
lexical H1 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory H1 2·6 — P1=1, P2=7 — G-aud 2 — Δ −P2=4

Stanza 13

13.1 ki-wa-si-ki-ya hu-ne-na + hu-ni-o-ngo-nga mo-yo-ni (8+8)
lexical = auditory 5·3 + 5·3 — P1=4, P2=7 | P1=4, P2=7 — G 2 + 2

13.2 sa-ru-fu ha-ku-na te-na + na-ha-u na-i-ta-ma-ni (8+8)
lexical = auditory 3·3·2 + 3·5 — P1=2, P2=5, P3=7 | P1=2, P2=7
— G 3 + 2 *reader says sarufu, not sarufi*

13.3 na ha-ta la-dha ha-ya-na + ka-ma ma-sha-pu ka-nwa-ni (8+8)
lexical = auditory 3·2·3 + 2?·3·3 — P1=2, P2=4, P3=7 | P1=1?, P2=4, P3=7
— G 3 + 3

13.4 si-e-le-wi hu-ne-na-ni + hu-i-mba a-u hu-tʰe-tʰa (8+8)
lexical = auditory 4·4 + 3·2?·3 — P1=3, P2=7 | P1=2, P2=4?, P3=7 — G 2 + 3

Stanza 14

14.1 la-u mu-ya-ka ta-ru-di + a-e te-na du-ni-ya-ni (8+8)
lexical = auditory 2·3·3 + 2?·2·4 — P1=1, P2=4, P3=7 | P1=1?, P2=3, P3=7
— G 3 + 3

14.2 mwa-na-ngu i-ta-m-bi-di + kwe-ne-nda ma-ha-ka-ma-ni (8+8)
lexical = auditory 3·5 + 3·5 — P1=2, P2=7 | P1=2, P2=7 — G 2 + 2

14.3 a-e-te na ma-sha-hi-di + wa-ni-yu-wa-o ya-ki-ni (8+8)
lexical = auditory 3·5 + 5·3 — P1=2, P2=7 | P1=4, P2=7 — G 2 + 2

14.4 nyu-tʰe mwe-nda ge-re-za-ni + kwa ha-ti-ya ku-wa-pa-ta (8+8)
lexical = auditory 2·2·4 + 4·4 — P1=1, P2=3, P3=7 | P1=3, P2=7 — G 3 + 2

Stanza 15

15.1 wa-lla-hi ha-mu-na ghe-ra + wa-la ha-mu-na i-ma-ni (8+8)
lexical = auditory 3·3?·2 + 2?·3?·3 — P1=2, P2=5?, P3=7 | P1=1?, P2=4?,
P3=7 — G 3 + 3 *Arabic realization of wallahi; no change of count*

15.2 ha-mu-na la ku-wa-ke-ra + ku-wa ha-mu-ni tha-mi-ni (8+8)
lexical = auditory 3·5 + 2?·3?·3 — P1=2, P2=7 | P1=1?, P2=4?, P3=7
— G 2 + 3

15.3 mi-mi ni ka-ma m-pwi-ra + hu-te-ze-wa u-wa-nda-ni (8+8)
lexical = auditory 2?·3?·3 + 4·4 — P1=1?, P2=4?, P3=7 | P1=3, P2=7
— G 3 + 2

15.4 hu-pi-jwa te-ke ndi-ya-ni + na ku-la a-na-e-pi-ta (8+8)
lexical = auditory 3?·2·3 + 3·5 — P1=2?, P2=4, P3=7 | P1=2, P2=7 — G 3 + 2

Stanza 16

16.1 ha-ta kwe-nye u-sha-i-ri + wa-so wa-ngu wa-me-bu-ni (8+8)
lexical = auditory 2?·2·4 + 2?·2·4 — P1=1?, P2=3, P3=7 | P1=1?, P2=3, P3=7
— G 3 + 3

16.2 zi-li-zo hu-ru ba-ha-ri + kwa ku-o-le-za wa-ge-ni (8+8)
lexical = auditory 3?·2·3 + 5·3 — P1=2?, P2=4, P3=7 | P1=4, P2=7 — G 3 + 2

16.3 mi-mi ha-yo si-ya-ki-ri + si ma-sha-i-ri ki-fa-ni (8+8)
lexical = auditory 2·2·4 + 5·3 — P1=1, P2=3, P3=7 | P1=4, P2=7 — G 3 + 2
16.4 ha-yo yo-te ni-kwa ni-ni + hi-zo ni mbi-nu za zi-tʰa (8+8)
lexical 2·2·2·2 + 2·3·3 — P1=1, P2=3, P3=5, P4=7 | P1=1, P2=4, P3=7
— G-lex 4 + 3
auditory 4·4 + 2·3·3 — P1=3, P2=7 | P1=1, P2=4, P3=7
— G-aud 2 + 3 — Δ −P1=1, −P3=5 | — *reader articulates aspirated tʰa, unnoticed in the first pass*

Stanza 17

17.1 ha-mbi-wa mwe-nye-we si-na + hi-ni ni a-ja-bu ga-ni (8+8)
lexical = auditory 3·3?·2 + 2·4?·2 — P1=2, P2=5?, P3=7 | P1=1, P2=5?, P3=7
— G 3 + 3
17.2 hu-wa-e ka-ko-sa shi-na + ka-wa na tʰa-ndu ya-nga-ni (8+8)
lexical = auditory 3·3?·2 + 2·3·3 — P1=2, P2=5?, P3=7 | P1=1, P2=4, P3=7
— G 3 + 3
17.3 nya-ni a-lo-ni-pa i-na + a-lo-na-ndi-ka ni nya-ni (8+8)
lexical = auditory 2·4?·2 + 5·3 — P1=1, P2=5?, P3=7 | P1=4, P2=7 — G 3 + 2
17.4 ki-wa si u-swa-hi-li-ni + ni wa-pi ni-li-po-pa-tʰa (8+8)
lexical = auditory 2·6 + 3·5 — P1=1, P2=7 | P1=2, P2=7 — G 2 + 2

Stanza 18

18.1 ku-wa we-ngi hu-ni-ne-na + si da-li-li a-si-la-ni (8+8)
lexical 2·2·4 + 4·4 — P1=1, P2=3, P3=7 | P1=3, P2=7 — G-lex 3 + 2
auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G-aud 2 + 2 — Δ −P1=1 | —
18.2 ya-ku-wa we-nye-we si-na + ki-nge-re-za ha-mu-o-ni (8+8)
lexical = auditory 3·3·2 + 4·4 — P1=2, P2=5, P3=7 | P1=3, P2=7 — G 3 + 2
18.3 hu-ne-nwa na we-ngi sa-na + pʰe-mbe zo-te du-ni-ya-ni (8+8)
lexical = auditory 3·3?·2? + 2?·2·4 — P1=2, P2=5?, P3=7? | P1=1?, P2=3,
P3=7 — G 3 + 3
18.4 ki-na na kwa-o shi-na-ni + mi-zi-ye ha-i-ku-kʰa-ta (8+8)
lexical = auditory 2?·3·3 + 3·5 — P1=1?, P2=4, P3=7 | P1=2, P2=7
— G 3 + 2 *miziye contracts mizizi yake, removing two syllables to fit the eight-syllable slot*

3.a. Divergences between the written text, the reading and the edition

The written text of 2005, published with a French translation in Roy (2013b, pp. 48–53), differs from the poet's reading in fourteen hemistiches: 3.4 H1 and H2, 5.2 H1, 5.3 H2, 6.1 H2, 6.3 H1, 7.4 H2, 9.4 H2, 10.1 H2, 11.3 H1, 12.4 H2, 14.3 H1, 14.4 H1 and 18.2 H2. In

each case the reading agrees with the version the poet published in 2013. Five of these divergences change the syllable count of the written text, which the reading restores to eight: *takapokuelea* (5.2 H1), *hazifia* (5.3 H2) and *Kiingereza* (18.2 H2) would give nine syllables, *Nabahany* (6.3 H1) and *mulipowata* (7.4 H2) seven. One changes the accentual grouping: the written *Ahe tena mashahidi* (14.3 H1) would yield three accentual groups, 2·2·4, where the reading *aete na mashahidi* yields two. Since the text studied here is the reading, none of these is a guni, a line defective in its count; they are listed so that a reader who compares the two versions can see where they part.

The reading and the version published in 2013 differ in seven hemistiches, given here with the reading first and the edition second: 4.4 H1 *mathumboni* / *mtumboni*; 5.3 H1 *zimebakiya* / *zimesaliya*; 7.2 H1 *mumenipuza* / *mumenipuuza*; 12.1 H2 *wa digrii* / *wa uzamili*; 13.1 H2 *huniongonga* / *huniungonga*; 13.2 H1 *sarufu* / *sarufi*; 18.1 H2 *asilani* / *aswilani*. In the first five, the written text of 2005 agrees with the reading, so that these are revisions the poet made for the edition; in the last two, it has a third spelling, *tharafu* and *athilani*. Two of them change the syllable count: at 7.2 the edition's spelling gives nine syllables where the poet reads eight, and at 12.1 it gives eight where the poet reads seven. Since the written text of 2005 also has *wa digrii*, 12.1 H2 counts as a guni of the text, and 7.2 H1 does not. The three versions also record aspiration differently: the written text marks it in seven words, the edition with an apostrophe in thirteen places, and the transcription as heard; they coincide only in part.

4. Statistical Analysis and Falsification

The mapping and the transitions are given in section 3. The pool comprises **141 hemistiches of 144**, three being excluded for false starts (1.4 H1, 10.4 H1, 12.4 H2); their counterpart hemistiches, verified independently, remain in the pool. Two further hemistiches are *guni* of the text, at 5.1 (H1, seven syllables) and 12.1 (H2, seven syllables), retained and counted against the matrix.

Condition 1. N = 8 holds in 139 of 141 hemistiches, that is 98.6 percent.

Condition 2 and Condition 4. The rhyme architecture is verified in section 1: local rhyme *a* on H1 of L1–L3, changing each stanza; global constant B (*-ni*) on H2 of L1–L3 and on H1 of L4; global constant X (*-ta*) on H2 of L4. All three hold in **18 stanzas of 18, without exception**, in their syllabic form. Aspiration, audible on the

rhyme syllable in four stanzas of the reading (1.4, 2.4, 16.4 and 17.4), is not treated as distinguishing a rhyme; the poet's own edition of 2013 writes *-ta* in all eighteen stanzas.

Condition 3. No word crosses the caesura in any of the seventy-two lines: seventy-two caesuras, seventy-two word boundaries, without exception.

Condition 5. G-lex = 2 in **79 hemistiches (56.0 percent)**, G-lex = 3 in 60 (42.6), G-lex = 4 in 2. G-aud = 2 in **86 (61.0 percent)**, G-aud = 3 in 54, G-aud = 4 in 1. The axiom is contradicted by 62 hemistiches on the lexical measure and 55 on the auditory one.

The transition. The reading departs from the lexical structure in **9 hemistiches of 141, that is 6.4 percent** — the lowest rate recorded for any *shairi* corpus in this chapter, and below the 12.5 percent of the prose control corpus of section 9.7.

Condition 6. The sequence 4·4 occurs in **29 hemistiches lexically (20.6 percent)** and **34 on the auditory measure (24.1 percent)**. Among the 79 hemistiches with G-lex = 2, the extended parity criterion (P_1 odd) holds in 37 (46.8 percent); among the 86 with G-aud = 2, in 45 (52.3 percent) — a near-even split, as in the *utenzi* corpora rather than the *wimbo*.

Condition 7. Nuclei of equal volume occur in **31 hemistiches of 141 (22.0 percent)**. Across consecutive hemistiches, the same sequence recurs in 16 pairs of 140 lexically (11.4 percent, against 17.8 expected under independence, 12.7 percent) and 18 of 140 auditorily (12.9 percent, against 19.2 expected, 13.7 percent) — in both cases below the independent expectation.

Condition 8. A tetrasyllabic nucleus occurs in **54 hemistiches (38.3 percent)**; only 29 are entirely partitioned into two such nuclei.

Inter-pass control. The two annotations, made under different notations at different times, coincide on **129 of 141 comparable hemistiches, that is 91.5 percent** — the highest rate in this book. Of the 12 divergent hemistiches, G falls from the first pass to the second in 5, rises in 2, and remains equal with a changed position in 5.

Verdict. The syllabic architecture — two global constants locked together, a *Grenzschema* real but at one hemistich in five, a parity split near even, no periodicity beyond chance — confirms, on the corpus most favorable to continuity with the classical tradition, what the *utenzi* and *wimbo* corpora already established.

A second, independent annotation of this poem exists: a full IPA transcription of the recitation, made by ear alone with the linguist

Marie-Françoise Rombi, prior to and independently of the present analysis. It is deposited alongside the recording (Roy & ibn Kadir, 2013) rather than tabulated here, for reasons of space; a reader wishing to compute an inter-judge coincidence rate may do so directly from the deposited material.

5. Conclusion on Kiswahili!!

This corpus is the one most favorable to the continuity thesis the accentual school would require. It is a modern classical *shairi*, composed by a poet reciting his own work, in the meter's strictest attested form: two global constants locked together across eighteen stanzas without exception, verified across two independent annotation passes at 91.5 percent agreement, the highest recorded in this study. If a bipartite accentual structure governs the *shairi* line, it should be most visible here.

A bipartite structure is present, and the fact should be stated without qualification. G = 2 governs 61.0 percent of hemistiches on the lexical measure and 61.7 percent on the auditory one. This is not a marginal presence. It is, however, contradicted in 55 hemistiches of 141 on the lexical measure and 54 on the auditory one, which places the axiom outside the reach of a generative law: a structure absent from more than a third of its instances describes a statistical tendency rather than a metrical rule.

The *Grenzschema* fares worse. The sequence 4·4 occurs in 20.6 percent of hemistiches lexically and 24.1 percent auditorily, a minority even among the hemistiches that already satisfy the bipartite count. The extended parity criterion, which admits three trochaic dispositions rather than one, yields a near-even split — 46.8 percent lexically, 52.3 percent auditorily — indistinguishable from the *utenzi* corpora and well short of what a generative principle would be expected to produce.

Periodicity is absent at both scales examined. Nuclei of equal volume, the precondition for any repeating foot, occur in 22.0 percent of hemistiches; in the remaining four cases out of five, no unit repeats, so that no foot can be extracted whatever is conceded regarding the count. Across consecutive hemistiches, identical sequences recur at a rate at or below what independent draws from the corpus distribution would produce — 11.4 against an expected 12.7 percent lexically, 12.9 against an expected 13.7 percent auditorily. The composer does not

sequence these configurations; the lexicon supplies them, and their distribution across lines is left undirected.

These results bear on the standing of Minerba's model in a manner distinct from the *Waji Waji*. There, the model failed for want of an object: the eight-syllable segment its axioms presuppose does not occur in a regular line of that poem. Here the segment is present in every hemistich, without exception, produced by a native speaker working in unbroken continuity with the classical tradition. Under these most favorable conditions, the model still accounts for a minority of what it predicts: a count correct in three hemistiches of five, a placement correct in barely half of those, a recoverable foot in one hemistich of five, and a line-to-line sequencing indistinguishable from chance.

The conjunction of these two outcomes is instructive. A corpus offering the model no object and a corpus offering it ideal conditions yield the same verdict: non-confirmation. A theory whose rate of confirmation does not improve when the evidential ground is made maximally favorable to it is not a theory contradicted by adverse data. What these two case studies together indicate is that no distribution of data, favorable or adverse, was positioned to discriminate the model from an arithmetical property of the Swahili lexicon operating within an eight-syllable frame.

9.6.3. Case Study VIII: Test of a Mashairi ya Vidato Matrix — *Uwongo*, by Muyaka bin Haji al-Ghassaniy (1776–1840)

The text is *Uwongo*, "Falsehood", transcribed on pages 92–93 of the *Diwani ya Muyaka* (Hichens, 1940). Muyaka is held to be the foundational figure of the classical Swahili tradition, and his structural templates were explicitly emulated in the twentieth century by poets such as Mathias E. Mnyampala.

The model tested here is the one formalized in section 2.5.2. Minerba derives the classical Swahili line from an imported tetrasyllabic template and states the derivation for the *utenzi*, the *shairi* and the *wimbo* alike (Minerba, 2025, p. 217). The model operates at two degrees of strictness, and both are tested: in its strict form it requires two tetrasyllabic nuclei with stresses on positions 3 and 7, which is the *Grenzschema*; in its extended form it requires only that the first stress fall on an odd syllable (Minerba, 2025, p. 220).

The syllabification and the accentual mapping follow the three-step protocol of section 9.2.3, restated in the preliminaries to this section. The accentual structure of each line is first derived from the printed text by the penultimate rule; the formal reading is then screened by ear, and the two results are reported side by side. Acoustic extraction is applied to a sample of hemistiches only, and it is reported for the three *shairi* corpora together in section 9.6.4.

1. Identification of the Formal Syllabic Matrix

The poem is a 16-syllable *shairi*, 8 + 8, with a stanzaic chaining device known as *Mashairi ya Vidato*. The rhyme is not confined to the stanza: it propagates across the whole poem.

Formal Parameters:

$n_{max} = 4$, $H = 2$, $S = 16$ (where $S = 8 + 8$).

The Matrix for Stanza k:

L_1: $[8, r_{k-1}] + [8, r_k]$ / /
L_2: $[8, r_{k-1}] + [8, r_k]$ / /
L_3: $[8, r_{k-1}] + [8, r_k]$ / /
L_4: $[8, r_k] + [8, X]$ / /

The Algebraic Equation for one stanza:

$$\boldsymbol{Shairi_{Vidato}(Stanza_k) = 3([8r_{k-1}] + [8r_k]//) \cdot ([8r_k] + [8X]//)}$$

Summed across the whole poem, for n stanzas:

$$\boldsymbol{Poem_{Vidato} = \sum_{k=1}^{n} [3([8r_{k-1}] + [8r_k]//) \cdot ([8r_k] + [8X]//)]}$$

The variable r_k is a rhyme that advances by one stanza at each step, each stanza's rhyme becoming the following stanza's local anchor. A global constant X closes L4 throughout the poem; in this composition, X is the syllable *-ngo*.

One further constraint governs the transition between stanzas in this poem and is not carried by the algebraic notation above: the first

hemistich of L1 in stanza k must repeat, verbatim, the first hemistich of L4 in stanza k − 1. The chaining is therefore lexical as well as prosodic, and it is verifiable directly against the mapping in section 3.

2. Uwongo: Syllabic and Accentual Mapping, Stanzas 1–4

Each line is given as two hemistiches separated by +, in the notation of the preliminaries to this section. Spelling follows the *Diwani ya Muyaka* as printed. Where the lexical and auditory readings coincide, a single line gives both; where they differ, the two are given separately, followed by Δ.

Stanza 1 — local rhyme r_0 = *-ma*, permuting rhyme r_1 = *-ye*, global constant *X* = *-ngo*

1.1 ka-ra-ta-si na-ku-tu-ma + le-o u-ni-tu-mi-ki-ye (8+8)
lexical = auditory 4·4 + 2·6 — P1=3, P2=7 | P1=1, P2=7 — G 2 + 2

1.2 ka-ba-le-ghe-she ka-li-ma + u-zi-di ku-se-ma na-ye (8+8)
lexical = auditory 5·3 + 3·3?·2 — P1=4, P2=7 | P1=2, P2=5?, P3=7 — G 2 + 3

1.3 ku-da-nga-nya-na si kwe-ma + ku-lla ja-mbo kwa ndi-a-ye (8+8)
lexical = auditory 5·3 + 2·2·4 — P1=4, P2=7 | P1=1, P2=3, P3=7 — G 2 + 3

1.4 ya-le-ke-ze ya-le-ke-ye + tu-sa-mbi-a-ne u-wo-ngo (8+8)
lexical = auditory 4·4 + 5·3 — P1=3, P2=7 | P1=4, P2=7 — G 2 + 2

tusiambiane is contracted to tusambiane to meet the eight-syllable count

Stanza 2 — local rhyme r_1 = *-ye*, permuting rhyme r_2 = *-ne*, global constant *X* = *-ngo*

2.1 ya-le-ke-ze ya-le-ke-ye + wa-o-le-zi wa-ya-o-ne (8+8)
lexical = auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G 2 + 2

2.2 tu-ji-vu-ne kha-ti-ma-ye + mi-ti-ma ye-tu i-so-ne (8+8)
lexical = auditory 4·4 + 3?·2·3 — P1=3, P2=7 | P1=2?, P2=4, P3=7 — G 2 + 3

2.3 ha-ku-na li-so ndu-u-ye + ma-mbo pi-a ya-fe-ne-ne (8+8)
lexical = auditory 3·2?·3 + 2?·2·4 — P1=2, P2=4?, P3=7 | P1=1?, P2=3, P3=7 — G 3 + 3

2.4 si-ta-ki tu-da-nga-nya-ne + na kwa-mbi-a-na u-wo-ngo (8+8)
lexical = auditory 3·5 + 5·3 — P1=2, P2=7 | P1=4, P2=7 — G 2 + 2

Stanza 3 — local rhyme r_2 = *-ne*, permuting rhyme r_3 = *-a*, global constant *X* = *-ngo*

3.1 si-ta-ki tu-da-nga-nya-ne + ndi-po ha-ya na-kwa-mbi-a (8+8)
lexical = auditory 3·5 + 2?·2·4? — P1=2, P2=7 | P1=1?, P2=3, P3=7? — G 2 + 3

3.2 la kwe-li tu-a-mbi-a-ne + zi-ta-ka-te ze-tu ni-a (8+8)
lexical = auditory 3·5 + 4·2·2 — P1=2, P2=7 | P1=3, P2=5, P3=7 — G 2 + 3
3.3 na u-fa-nya-po ma-ngi-ne + mwi-sho wa-kwe ni vi-ba-ya (8+8)
lexical = auditory 5·3 + 2?·2·4 — P1=4, P2=7 | P1=1?, P2=3, P3=7 — G 2 + 3
3.4 vi-ta-ku-ja ku-tu-ki-a + tu-sa-mbi-a-ne u-wo-ngo (8+8)
lexical = auditory 4·4 + 5·3 — P1=3, P2=7 | P1=4, P2=7 — G 2 + 2
tusiambiane again contracted to tusambiane
Stanza 4 — local rhyme r_3 = *-a*, permuting rhyme r_4 = *-zi*, global constant *X* = *-ngo*
4.1 vi-ta-ku-ja ku-tu-ki-a + u-ya-o-ne ma-tu-ku-zi (8+8)
lexical = auditory 4·4 + 4·4 — P1=3, P2=7 | P1=3, P2=7 — G 2 + 2
4.2 na kwa-mba wa-pe-nde-le-a + ne-le-za-to wa-zi-wa-zi (8+8)
lexical 3·5 + 4·4 — P1=2, P2=7 | P1=3, P2=7 — G-lex 2 + 2
auditory 3·5 + 4·2·2 — P1=2, P2=7 | P1=3, P2=5, P3=7
— G-aud 2 + 3 — Δ — | +P2=5
4.3 ni-pa-te ya-fa-ha-mi-a + ni-fu-nu-e ma-si-ki-zi (8+8)
lexical = auditory 3·5 + 4·4 — P1=2, P2=7 | P1=3, P2=7 — G 2 + 2
4.4 tu-u-wa-te u-pu-u-zi + tu-sa-mbi-a-ne u-wo-ngo (8+8)
lexical = auditory 4·4 + 5·3 — P1=3, P2=7 | P1=4, P2=7 — G 2 + 2
tusiambiane again contracted to tusambiane

3. Statistical Analysis and Falsification

The mapping is given in section 2. The corpus comprises four stanzas of four lines, that is thirty-two hemistiches — the smallest corpus of this chapter, and the results below should be weighed accordingly. Its interest lies less in the size of the sample than in the demands the *vidato* architecture places on it, and in the fact, established below, that this is the corpus most favorable to the accentual model in the whole of this study.

3.a. Conditions 1, 2, 3 and 4: the syllabic constant, the caesura, and the stanzaic chain

Condition 1. N = 8 holds in all thirty-two hemistiches, without a single *guni* of the text.

Condition 2. The rhyme chain is initialized by r_0 (*-ma*), which governs H1 of stanza 1 alone and does not itself descend from an earlier stanza. From r_1 onward, each permuting rhyme advances by one stanza at each step — r_1 (*-ye*), r_2 (*-ne*), r_3 (*-a*) — governing H2 of L1–L3 in the stanza where it is established and, by the *vidato*

transition, H1 of L1–L3 in the stanza that follows, in accordance with the matrix of section 1. Four of the five rhyme chains hold in their strict syllabic form. The fifth, r_3 (*-a*), is contradicted once: at 3.3, *vibaya* closes on the vowel alone where *nia* and the chain otherwise close on the full syllable. X^s is tested first, as the hierarchy requires; it fails on this single line; X^v is tested in turn and holds. The global constant *X*, *-ngo*, closes every fourth line of every stanza without exception, in its syllabic form.

Condition 3. No word crosses the caesura in any of the sixteen lines: sixteen caesuras, sixteen word boundaries, without exception.

Condition 4. The *vidato* architecture is tested here in its formal sense: whether the rhyme permutation itself, as the recursive equation of section 1 specifies, propagates without a broken link. The permuting rhyme r_k, which closes H2 of L1–L3 within its own stanza, becomes the local rhyme governing H1 of L1–L3 in the stanza that follows — r_1(-ye) descending into stanza 2, r_2(-ne) into stanza 3, r_3(-a) into stanza 4 — and this holds at every one of the three transitions without exception. For a poem of $n = 4$ stanzas, the summative equation,

$$\mathbf{Poem_{Vidato}} = \sum_{k=1}^{4} [3([8r_{k-1}] + [8r_k]//) \cdot ([8r_k] + [8X]//)]$$

predicts four instantiations of the stanza matrix chained in this way and closed, in every fourth line, by the constant *X*. The mapping confirms all four instantiations directly, term by term: nothing in the equation is asserted without a corresponding line in the text that realizes it.

A further constraint governs this poem beyond what the algebraic matrix formalizes, and it is reported separately rather than folded into Condition 4: the opening hemistich of each stanza repeats, verbatim, the closing hemistich of the stanza before it. The three transitions of this poem — stanza 1 to 2, 2 to 3, 3 to 4 — all satisfy this lexical rule without exception. *Yalekeze yalekeye* closes stanza 1 and opens stanza 2; *Sitaki tudanganyane* closes stanza 2 and opens stanza 3; *Vitakuja kutukia* closes stanza 3 and opens stanza 4. This repetition and the rhyme's descent operate at the same textual boundary but are not the same fact: one is a property of the whole hemistich's wording, the other of its syllable's rhyme class alone, and the falsification test of Condition 4 bears only on the second.

3.b. Condition 8: the tetrasyllabic foot, arithmetically

Since both hemistiches measure eight syllables, a multiple of four, the tetrasyllabic template is arithmetically available throughout this corpus, unlike the *Waji Waji* and the hyper-fractionated *wimbo*. The condition is not decisive here on arithmetical grounds alone; what it can settle is whether the template, being available, generalizes — which is the question the remaining conditions address.

3.c. Condition 5: the bipartite accentual axiom

The lexicon yields G = 2 in **24 of 32 hemistiches, that is 75.0 percent**, and G = 3 in 8, that is 25.0 percent. The reading realizes G = 2 in **23, that is 71.9 percent**, and G = 3 in 9, that is 28.1 percent.

These are the highest figures for the bipartite count recorded anywhere in this book. Three of every four hemistiches of *Uwongo* carry exactly two accentual groups, on the lexical measure. The result should be stated as such and not qualified away: on this corpus, composed by the figure the tradition holds as its master, the bipartite tendency is at its strongest.

It remains a tendency and not an axiom. Eight hemistiches on the lexical measure and nine on the auditory one contradict it, and an axiom that tolerates one exception in four has already ceased to function as one.

The transition. The reading departs from the lexical structure in **1 hemistich of 32, that is 3.1 percent** — the lowest rate of any corpus in this chapter, below the 6.4 percent of *Kiswahili!!* and well below the 12.5 percent of the prose control corpus of section 9.7. The single departure is reported in the mapping: at 4.2, the reading adds a nucleus, 4·2·2 for a lexical 4·4.

3.d. Condition 6: the *Grenzschema*

The sequence 4·4 occurs in **11 hemistiches of 32 on the lexical measure, that is 34.4 percent**, and in **10, that is 31.2 percent**, on the auditory one. These are, again, the highest figures recorded for this configuration in the book — nearly double the rate of the *Utendi wa Haudaji* and exceeding even the *Rasi 'lGhuli*.

The figure must be read against what it is not. A configuration present in a little over a third of the corpus is not a generative core: it is the modal outcome among several, in a poem where the composer's own lexicon happens, more often than in any other corpus examined, to supply two words of four syllables. Among the twenty-four

hemistiches that satisfy the bipartite axiom on the lexical measure, 4·4 accounts for eleven, that is 45.8 percent — a plurality, not a majority, and one that a poem two-thirds its size could as easily have failed to produce.

The extended parity criterion. Among the twenty-four hemistiches with G-lex = 2, the first stress falls on an odd syllable in 12, that is exactly 50.0 percent; among the twenty-three with G-aud = 2, in 11, that is 47.8 percent. Chance itself could not divide the corpus more evenly.

3.e. Condition 7: periodicity

Nuclei of equal volume occur in **11 hemistiches of 32, that is 34.4 percent** — again the highest rate of internal symmetry recorded in this book, and still a minority. In the remaining twenty-one, no unit repeats within the hemistich.

Across hemistiches, the same sequence recurs in 5 pairs of 31 on the lexical measure, that is 16.1 percent, against 6.6 expected under independence, that is 21.3 percent; on the auditory measure, in 5 of 31, that is 16.1 percent, against 6.1 expected, that is 19.5 percent. On both measures the observed rate of recurrence falls below what independent draws from the corpus distribution would produce. If Muyaka sequenced these configurations deliberately, the sequencing does not appear as repetition; if anything, it appears as its avoidance.

3.f. Inter-pass control

The two annotations coincide on **30 of 32 hemistiches, that is 93.8 percent** — the highest agreement recorded for any corpus of this chapter after *Kiswahili!!*. Of the two divergent hemistiches, one records an additional accentual group in the second pass and one records the same number with a changed position.

3.g. What this corpus establishes

Uwongo is not a corpus that struggles to meet the terms of the accentual model, as the *Waji Waji* did by arithmetic, or that meets them only in part, as *Kiswahili!!* did. It is the corpus, among all eight examined in this study, most favorable to every measure the model proposes: the highest rate of G = 2, the highest rate of the *Grenzschema*, the highest rate of internal symmetry, the lowest rate of departure between lexicon and reading, the highest inter-pass agreement. If native mastery of the classical *shairi* inclined a

composer toward the configurations the accentual school describes, this is where that inclination should be most visible.

It is visible, to a degree, and the degree is the finding. Three hemistiches in four carry two accents; one in three carries the *Grenzschema*; one in three exhibits internal symmetry; parity divides at a coin's toss. Each of these figures is the ceiling this study has recorded, and each remains short of what a generative rule requires. The composer most steeped in the tradition does not approach the trochaic engine more closely because he reaches for it; he approaches it because the density of tetrasyllabic and disyllabic words in his lexicon is itself denser than elsewhere in this corpus, for reasons the accentual model does not examine and the syllabic one does not need to.

The *vidato* architecture, verified in Condition 4, bears on this differently and should be kept apart from the accentual measures above rather than folded into them. It operates through two mechanisms locked together — a hemistich repeated verbatim across a stanza boundary, and a rhyme that descends unchanged from the terminus of three lines into the opening of the next stanza's own three lines — and both are facts legible on the printed page, requiring no listening whatever. A model that locates the meter in an accentual pulse has no means of registering either, since neither depends on where a stress falls. The full articulation of this architecture, and what it establishes against the exogenous model, is set out in the conclusion below.

4. Conclusion on Uwongo

This poem was selected because it offers the accentual school its most favorable ground: a text by the figure the tradition itself holds as the master of the classical *shairi*, in a corpus small enough that no distribution could be attributed to noise, and constructed under a stanzaic architecture, *Mashairi ya Vidato*, whose demands on the composer exceed those of ordinary chaining. The results bear this out on every measure the model proposes, and they should be reported for what they are before anything further is drawn from them.

Three hemistiches in four carry exactly two accentual groups, the highest proportion recorded anywhere in this study. One hemistich in three exhibits the *Grenzschema*, again the highest recorded rate, nearly double that of the *Utendi wa Haudaji*. Internal symmetry, the precondition for any repeating foot, reaches its own ceiling here at one

hemistich in three. The reading departs from the lexical structure in one hemistich in thirty-two, the lowest transition rate in the book. On every axis this study has measured, *Uwongo* is the corpus that comes closest to what the accentual model predicts.

It does not reach it. A rule holding in three cases of four is contradicted in the fourth; a prototype present in a third of the corpus is absent from the other two-thirds; symmetry recorded once in three hemistiches leaves no unit to repeat in the remaining two. The extended parity criterion, tested on the lexicon and on the reading alike, divides the corpus within a percentage point of an even split — the outcome chance alone would produce. And the sequence of configurations from one hemistich to the next recurs at a rate below what independent draws from the corpus would predict, on both measures. A composer working, by hypothesis, toward a trochaic pulse he intends to sustain does not produce a sequence indistinguishable from one that avoids repetition.

What these figures indicate is not that Muyaka approximates the accentual model more closely than other poets. It is that his lexicon, in this particular composition, happens to supply proportionally more words of two and four syllables than the lexicons of the other corpora examined in this chapter, and that a model measuring nothing but the arithmetic of word length inside an eight-syllable frame will register that density as apparent conformity, wherever it occurs. The ceiling reached here is not evidence of an approach to a generative rule; it is the clearest illustration this study affords of what such a model actually measures.

The *Mashairi ya Vidato* architecture is a separate matter, and it carries a weight in this study that the accentual reckoning above does not. Condition 4 verifies the rhyme propagation the recursive equation predicts: the permuting rhyme that closes the second hemistich of a stanza's first three lines descends, unchanged, to govern the first hemistich of those same three lines in the stanza that follows — r_1 into stanza 2, r_2 into stanza 3, r_3 into stanza 4 — confirmed at every one of the three transitions.

The summative equation of section 1, states this in closed form for the poem's four stanzas:

$$\text{Poem}_{\text{Vidato}} = \sum_{k=1}^{4} [\mathbf{3}([\mathbf{8}\boldsymbol{r}_{k-1}] + [\mathbf{8}\boldsymbol{r}_{k}]//) \cdot ([\mathbf{8}\boldsymbol{r}_{k}] + [\mathbf{8}\boldsymbol{X}]//)]$$

The mapping confirms every term of it directly against the printed text: nothing in the equation is asserted without a line that realizes it.

A further constraint, additional to this formal architecture and not carried by the algebraic notation, governs the same textual boundary: the opening hemistich of each stanza repeats, verbatim, the closing hemistich of the stanza before it, three times over without exception. The two facts sit at the same seam in the text but are not the same claim — one bears on the wording of the whole hemistich, the other on the rhyme class of its terminal syllable alone — and only the second is what Condition 4 formally tests.

None of this is measured by counting accents, because nothing in the chain depends on where a stress falls. A model built to locate the meter in a beat has no means of registering an architecture that operates on the systemic propagation of a rhyme chain across stanza boundaries. The four stanzas of *Uwongo* are bound into a single continuous structure by this precise recursive mechanism—one that is fully apparent to a model formalizing syllabic and terminal constraints, but unavailable, in principle rather than by omission, to one that listens for a pulse.

9.6.4. Instrumental Verification of the Hostile Sample: *Waji Waji*, *Kiswahili!!*, and *Uwongo* Combined

Instrumental analysis is applied to a sample of hemistiches drawn across the three *shairi* corpora of this section, in accordance with the protocol of section 9.2.3. Exhaustive verification of these corpora is auditory; no statistic reported in sections 9.6.1 to 9.6.3 rests on the measurements below.

1. Composition of the sample

Eleven hemistiches are measured, and the composition was fixed before any measurement was made, on three principles.

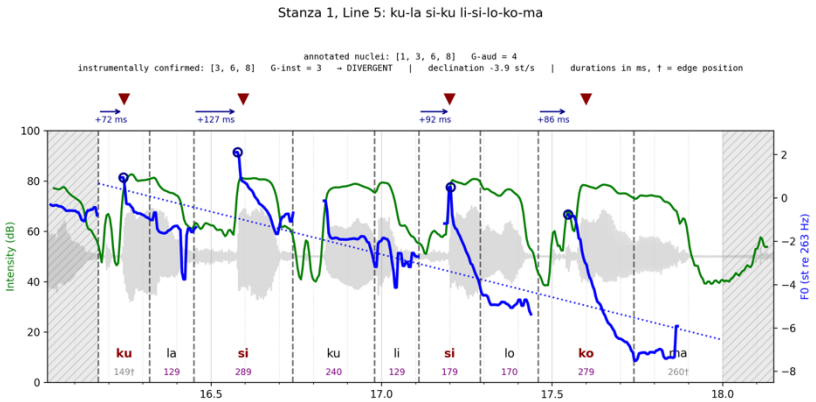
Six by reasoned selection, against the model rather than for it. Two per corpus, chosen by two criteria stated in advance: the highest G attested in that corpus, against the bipartite axiom; and, among hemistiches with G = 2, the greatest asymmetry between nuclei, against the *Grenzschema*. The corpus most favorable to the model, *Uwongo*, is not exempted from this scrutiny; its own extremes are tested like the others'.

Two by random draw, on the pooled population of 301 retained hemistiches across the three corpora, seeded as throughout this study, the six reasoned selections excluded from the draw.

Three for the *Grenzschema* itself, one per corpus, the first unmarked instance of 4·4 encountered in the text of each — since a sample built only to attack the model would not be a fair test of it, and the ideal configuration is measured on the same footing as its contradictions.

2. Figures

Figure 1. *Waji Waji*, stanza 1, line 5, second hemistich: *kula siku lisilokoma*.



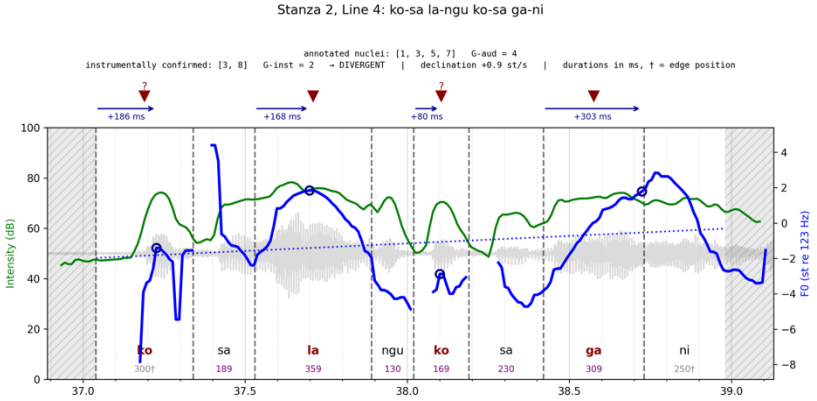
Waji Waji, Stanza 1, Line 5, H2: "ku-la si-ku li-si-lo-ko-ma"
Lexical reading: 2·2·3·2, P_1=1, P_2=3, P_3=6, P_4=8, G-lex = 4.
Auditory reading: 2·2·3·2, P_1=1, P_2=3, P_3=6, P_4=8, G-aud = 4, delta –
Instrumental reading: 4·3·2, P_1=3, P_2=6, P_3=8, G-inst = 3, delta –P_1=1 (against the lexical reading).
Speaker median 263 Hz; declination –3.94 st/s.

The instrumental reading diverges from the auditory reading: position(s) [1] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	ku	149 ms†	F0 -0.66 st	peak + 72 ms	intensity	69.7 dB
2	la	129 ms	F0 -1.45 st	peak + 17 ms	intensity	72.7 dB
3	si	289 ms	F0 +0.92 st	peak +127 ms	intensity	69.2 dB
4	ku	240 ms	F0 +0.45 st	peak + 95 ms	intensity	66.0 dB
5	li	129 ms	F0 +0.08 st	peak + 28 ms	intensity	68.7 dB
6	si	179 ms	F0 +1.48 st	peak + 92 ms	intensity	68.7 dB
7	lo	170 ms	F0 -0.80 st	peak +134 ms	intensity	69.4 dB
8	ko	279 ms	F0 +0.69 st	peak + 86 ms	intensity	67.4 dB
9	ma	260 ms†	F0 -1.36 st	peak +128 ms	intensity	60.2 dB

Figure 2. *Kiswahili!!*, stanza 2, line 4, first hemistich: *kosa langu kosa gani*.



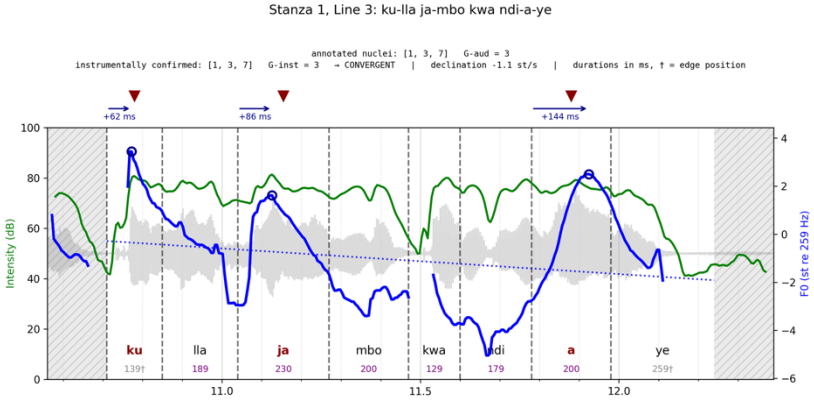
Kiswahili!!, Stanza 2, Line 4, H1: "ko-sa la-ngu ko-sa ga-ni"
Lexical reading: 2·2·2·2, P_1=1, P_2=3, P_3=5, P_4=7, G-lex = 4.
Auditory reading: 2?·2·2?·2, P_1=1?, P_2=3, P_3=5?, P_4=7, G-aud = 4, delta —.
Instrumental reading: 4·4, P_1=3, P_2=8, G-inst = 2, delta $-P_1$=1, $-P_3$=5, $-P_4$=7, $+P_2$=8 (against the lexical reading).
Speaker median 123 Hz; declination +0.85 st/s.

The instrumental reading diverges from the auditory reading: position(s) [1, 5, 7] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [8] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

Per-syllable measurements:

1	ko	300 ms†	F0 -1.08 st	peak +186 ms	intensity	58.2 dB
2	sa	189 ms	F0 +1.31 st	peak + 57 ms	intensity	64.1 dB
3	la	359 ms	F0 +1.79 st	peak +168 ms	intensity	73.8 dB
4	ngu	130 ms	F0 -2.86 st	peak + 9 ms	intensity	64.9 dB
5	ko	169 ms	F0 -2.44 st	peak + 80 ms	intensity	61.1 dB
6	sa	230 ms	F0 -3.10 st	peak + 91 ms	intensity	59.5 dB
7	ga	309 ms	F0 +0.25 st	peak +303 ms	intensity	71.1 dB
8	ni	250 ms†	F0 +1.29 st	peak + 33 ms	intensity	70.1 dB

Figure 3. *Uwongo*, stanza 1, line 3, second hemistich: *kulla jambo kwa ndiaye.*

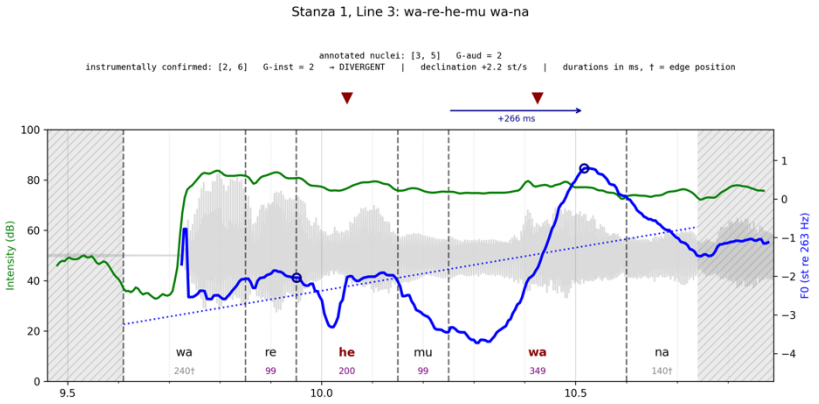



```
Uwongo, Stanza 1, Line 3, H2: "ku-lla ja-mbo kwa ndi-a-ye"
Lexical reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-lex = 3.
Auditory reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-aud = 3, delta —.
Instrumental reading: 2·2·4, P₁=1, P₂=3, P₃=7, G-inst = 3, delta —
(against the lexical reading).
Speaker median 259 Hz; declination –1.07 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 ku     139 ms†  F0 +2.40 st  peak + 62 ms  intensity  69.7 dB
  2 lla    189 ms   F0 -0.05 st  peak +  3 ms  intensity  74.9 dB
  3 ja     230 ms   F0 +0.60 st  peak + 86 ms  intensity  75.3 dB
  4 mbo    200 ms   F0 -1.63 st  peak +  4 ms  intensity  71.9 dB
  5 kwa    129 ms   F0 -1.92 st  peak + 62 ms  intensity  68.0 dB
  6 ndi    179 ms   F0 -2.52 st  peak +176 ms  intensity  74.1 dB
  7 a      200 ms   F0 +2.06 st  peak +144 ms  intensity  76.1 dB
  8 ye     259 ms†  F0 +1.11 st  peak +  1 ms  intensity  58.8 dB
```

Figure 4. *Waji Waji*, stanza 1, line 3, first hemistich: *warehemu wana*.

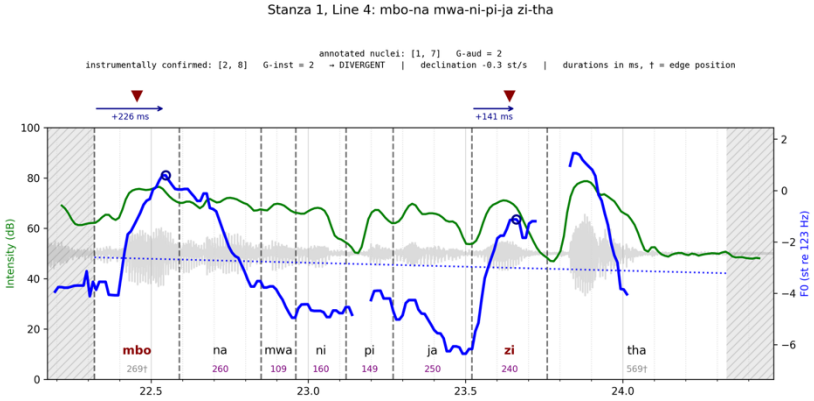



```
Waji Waji, Stanza 1, Line 3, H1: "wa-re-he-mu wa-na"
Lexical reading: 4·2, P₁=3, P₂=5, G-lex = 2.
Auditory reading: 4·2, P₁=3, P₂=5, G-aud = 2, delta —.
Instrumental reading: 3·3, P₁=2, P₂=6, G-inst = 2, delta −P₁=3, −P₂=5,
+P₁=2, +P₂=6 (against the lexical reading).
Speaker median 263 Hz; declination +2.23 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [3, 5] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2, 6] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 wa     240 ms†  F0 +0.56 st  peak +119 ms  intensity  59.9 dB
  2 re      99 ms   F0 +0.58 st  peak + 55 ms  intensity  81.2 dB
  3 he     200 ms   F0 -0.03 st  peak +  0 ms  intensity  77.9 dB
  4 mu      99 ms   F0 -1.01 st  peak +  0 ms  intensity  75.7 dB
  5 wa     349 ms   F0 -0.11 st  peak +266 ms  intensity  76.1 dB
  6 na     140 ms†  F0 +0.15 st  peak +  2 ms  intensity  74.5 dB
```

Figure 5. *Kiswahili!!*, stanza 1, line 4, second hemistich: *mbona mwanipija zitha*.

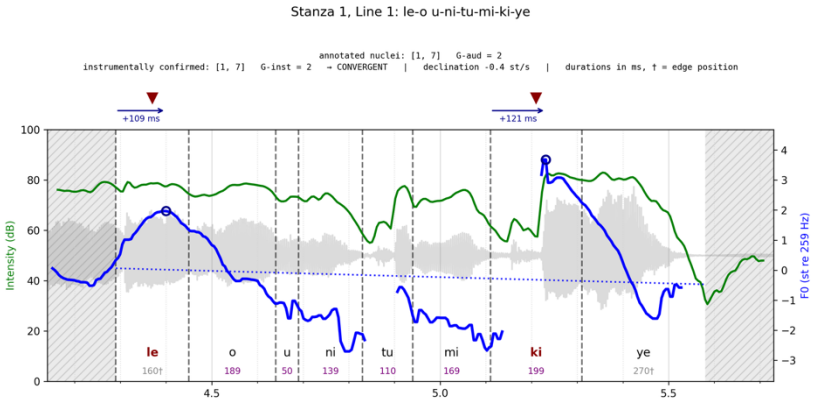


Kiswahili!!, Stanza 1, Line 4, H2: "mbo-na mwa-ni-pi-ja zi-tha"
Lexical reading: 2·4·2, P_1=1, P_2=5, P_3=7, G-lex = 3.
Auditory reading: 2·6, P_1=1, P_2=7, G-aud = 2, delta $-P_2$=5.
Instrumental reading: 3·5, P_1=2, P_2=8, G-inst = 2, delta $-P_1$=1, $-P_2$=5, $-P_3$=7, $+P_1$=2, $+P_2$=8 (against the lexical reading).
Speaker median 123 Hz; declination −0.31 st/s.

The instrumental reading diverges from the auditory reading: position(s) [1, 7] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2, 8] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 mbo     269 ms†  F0 +1.09 st  peak +226 ms  intensity  70.8 dB
  2 na      260 ms   F0 +1.12 st  peak + 26 ms  intensity  70.0 dB
  3 mwa     109 ms   F0 -1.43 st  peak + 37 ms  intensity  68.2 dB
  4 ni      160 ms   F0 -1.84 st  peak + 38 ms  intensity  63.4 dB
  5 pi      149 ms   F0 -1.39 st  peak +109 ms  intensity  59.8 dB
  6 ja      250 ms   F0 -2.42 st  peak + 70 ms  intensity  63.6 dB
  7 zi      240 ms   F0 +0.56 st  peak +141 ms  intensity  61.9 dB
  8 tha     569 ms†  F0 +2.46 st  peak + 92 ms  intensity  59.0 dB
```

Figure 6. *Uwongo*, stanza 1, line 1, second hemistich: *leo unitumikiye*.

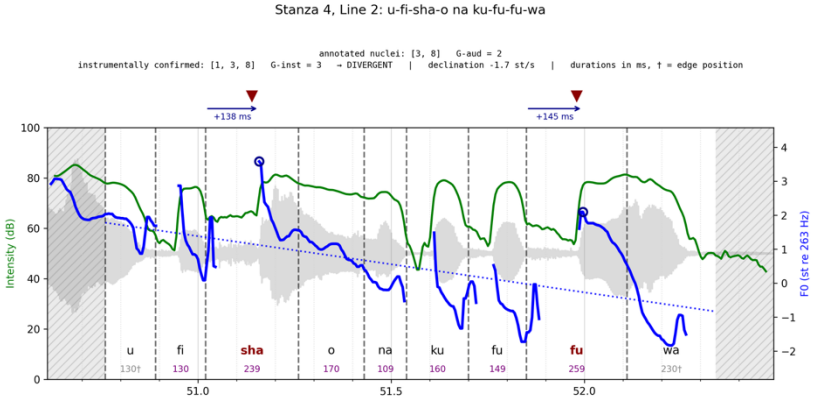



```
Uwongo, Stanza 1, Line 1, H2: "le-o u-ni-tu-mi-ki-ye"
Lexical reading: 2·6, P₁=1, P₂=7, G-lex = 2.
Auditory reading: 2·6, P₁=1, P₂=7, G-aud = 2, delta —.
Instrumental reading: 2·6, P₁=1, P₂=7, G-inst = 2, delta — (against
the lexical reading).
Speaker median 259 Hz; declination –0.42 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 le     160 ms†  F0 +1.48 st  peak +109 ms  intensity  77.5 dB
  2 o      189 ms   F0 +0.16 st  peak +  2 ms  intensity  76.1 dB
  3 u       50 ms   F0 -1.11 st  peak + 41 ms  intensity  72.5 dB
  4 ni     139 ms   F0 -1.71 st  peak + 82 ms  intensity  70.4 dB
  5 tu     110 ms   F0 -0.83 st  peak + 85 ms  intensity  65.7 dB
  6 mi     169 ms   F0 -1.62 st  peak + 18 ms  intensity  71.5 dB
  7 ki     199 ms   F0 +2.01 st  peak +121 ms  intensity  69.5 dB
  8 ye     270 ms†  F0 +0.45 st  peak +  2 ms  intensity  69.1 dB
```

Figure 7 (random draw). *Waji Waji*, stanza 4, line 2, second hemistich: *ufishao na kufufuwa*.

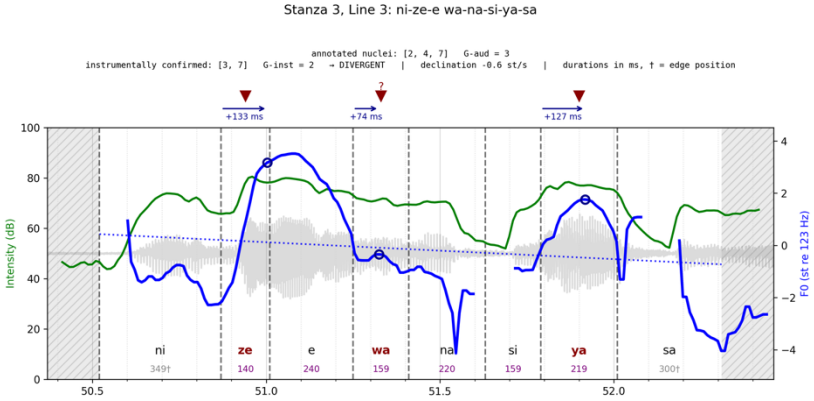



```
Waji Waji, Stanza 4, Line 2, H2: "u-fi-sha-o na ku-fu-fu-wa"
Lexical reading: 4·5, P₁=3, P₂=8, G-lex = 2.
Auditory reading: 4·5, P₁=3, P₂=8, G-aud = 2, delta —.
Instrumental reading: 2·2·5, P₁=1, P₂=3, P₃=8, G-inst = 3, delta +P₁=1
(against the lexical reading).
Speaker median 263 Hz; declination –1.66 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [1] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 u      130 ms†  F0 +0.02 st  peak +113 ms  intensity  71.7 dB
  2 fi     130 ms   F0 -0.27 st  peak + 64 ms  intensity  65.3 dB
  3 sha    239 ms   F0 +0.60 st  peak +138 ms  intensity  71.1 dB
  4 o      170 ms   F0 +0.19 st  peak +  2 ms  intensity  75.2 dB
  5 na     109 ms   F0 -0.59 st  peak +  0 ms  intensity  71.2 dB
  6 ku     160 ms   F0 -0.77 st  peak + 71 ms  intensity  64.4 dB
  7 fu     149 ms   F0 -0.75 st  peak + 65 ms  intensity  67.5 dB
  8 fu     259 ms   F0 +1.22 st  peak +145 ms  intensity  68.3 dB
  9 wa     230 ms†  F0 -0.49 st  peak +  3 ms  intensity  67.8 dB
```

Figure 8 (random draw). *Kiswahili!!*, stanza 3, line 3, first hemistich: *nizee wanasiasa*.

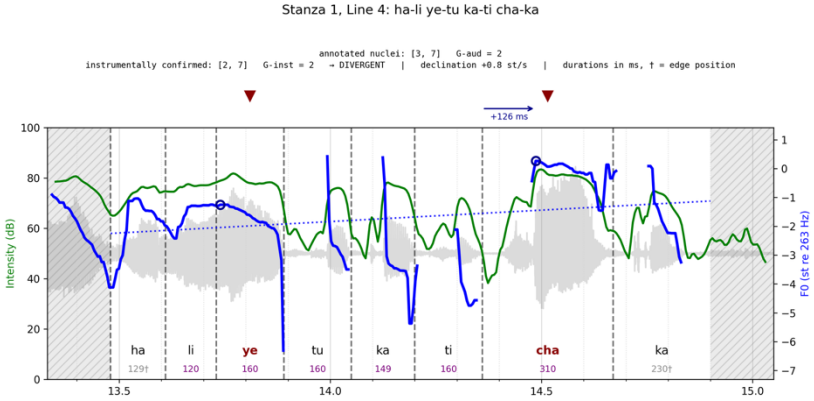



```
Kiswahili!!, Stanza 3, Line 3, H1: "ni-ze-e wa-na-si-ya-sa"
Lexical reading: 3·5, P₁=2, P₂=7, G-lex = 2.
Auditory reading: 3·2?·3, P₁=2, P₂=4?, P₃=7, G-aud = 3, delta +P₂=4.
Instrumental reading: 4·4, P₁=3, P₂=7, G-inst = 2, delta −P₁=2, +P₁=3
(against the lexical reading).
Speaker median 123 Hz; declination −0.65 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [2, 4] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [3] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 ni      349 ms†  F0 -1.56 st  peak + 81 ms  intensity  63.7 dB
  2 ze      140 ms   F0 +0.50 st  peak +133 ms  intensity  73.8 dB
  3 e       240 ms   F0 +2.58 st  peak + 73 ms  intensity  76.9 dB
  4 wa      159 ms   F0 -0.51 st  peak + 74 ms  intensity  70.6 dB
  5 na      220 ms   F0 -1.41 st  peak + 35 ms  intensity  65.0 dB
  6 si      159 ms   F0 -0.44 st  peak +156 ms  intensity  61.1 dB
  7 ya      219 ms   F0 +1.55 st  peak +127 ms  intensity  76.4 dB
  8 sa      300 ms†  F0 -1.12 st  peak + 68 ms  intensity  63.0 dB
```

Figure 9 (*Grenzschema*). *Waji Waji*, stanza 1, line 4, second hemistich: *hali yetu kati chaka*.

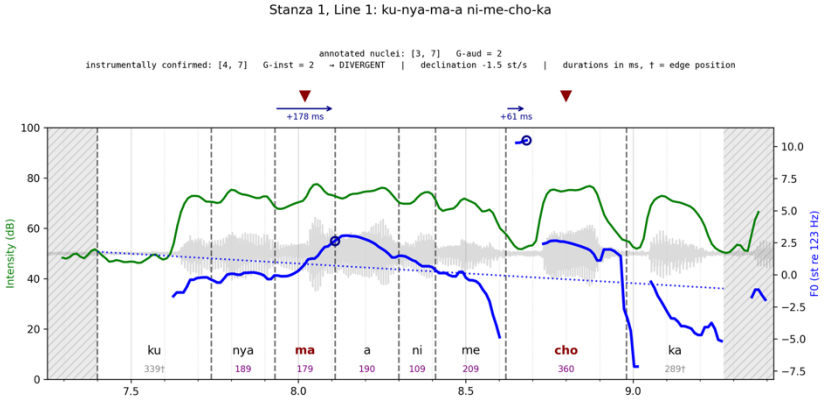



```
Waji Waji, Stanza 1, Line 4, H2: "ha-li ye-tu ka-ti cha-ka"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4·4, P₁=3, P₂=7, G-aud = 2, delta —.
Instrumental reading: 3·5, P₁=2, P₂=7, G-inst = 2, delta −P₁=3, +P₁=2
(against the lexical reading).
Speaker median 263 Hz; declination +0.79 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [3] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [2] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3. Fundamental frequency is one of three parameters and is not decisive on its own; the durations and intensities given below are reported whether or not they agree with it.

```
Per-syllable measurements:
  1 ha     129 ms†  F0 +0.19 st  peak + 60 ms  intensity  72.5 dB
  2 li     120 ms   F0 +0.48 st  peak +106 ms  intensity  76.5 dB
  3 ye     160 ms   F0 +0.16 st  peak +  9 ms  intensity  78.5 dB
  4 tu     160 ms   F0 -0.77 st  peak +103 ms  intensity  62.9 dB
  5 ka     149 ms   F0 -1.65 st  peak + 74 ms  intensity  66.2 dB
  6 ti     160 ms   F0 -2.31 st  peak + 96 ms  intensity  60.1 dB
  7 cha    310 ms   F0 +1.23 st  peak +126 ms  intensity  69.9 dB
  8 ka     230 ms†  F0 -0.35 st  peak + 83 ms  intensity  60.6 dB
```

Figure 10 (*Grenzschema*). *Kiswahili!!*, stanza 1, line 1, first hemistich: *kunyamaa nimechoka*.

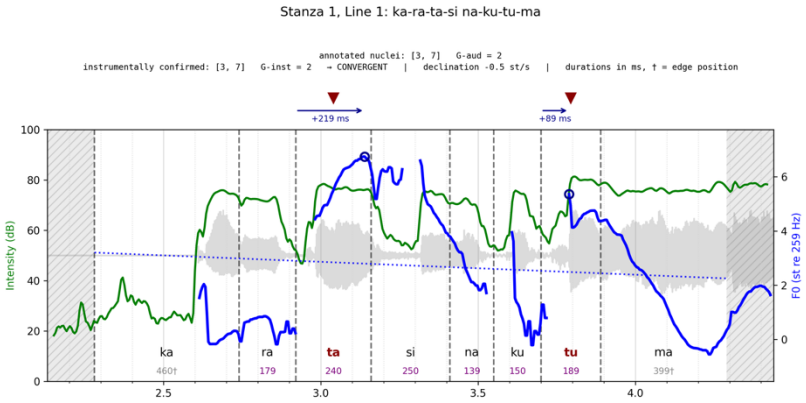



```
Kiswahili!!, Stanza 1, Line 1, H1: "ku-nya-ma-a ni-me-cho-ka"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4·4, P₁=3, P₂=7, G-aud = 2, delta —.
Instrumental reading: 5·3, P₁=4, P₂=7, G-inst = 2, delta −P₁=3, +P₁=4
Speaker median 123 Hz; declination −1.53 st/s.
```

The instrumental reading diverges from the auditory reading: position(s) [3] annotated by ear are not designated by the instrument, their detrended F0 mean being either below the declination or lower than that of an adjacent syllable; and position(s) [4] are designated by the instrument without having been annotated by ear. The divergence is recorded and the annotation is not revised, in accordance with the rule of non-revision stated in section 9.2.3.

```
Per-syllable measurements:
  1 ku      339 ms†  F0 -2.09 st  peak +327 ms  intensity  54.9 dB
  2 nya     189 ms   F0 -1.22 st  peak +127 ms  intensity  72.4 dB
  3 ma      179 ms   F0 +0.15 st  peak +178 ms  intensity  72.2 dB
  4 a       190 ms   F0 +1.87 st  peak + 39 ms  intensity  73.9 dB
  5 ni      109 ms   F0 +0.65 st  peak +  9 ms  intensity  71.8 dB
  6 me      209 ms   F0 -1.00 st  peak + 80 ms  intensity  67.5 dB
  7 cho     360 ms   F0 +3.48 st  peak + 61 ms  intensity  65.6 dB
  8 ka      289 ms†  F0 -2.98 st  peak + 73 ms  intensity  63.1 dB
```

Figure 11 (*Grenzschema*). *Uwongo*, stanza 1, line 1, first hemistich: *karatasi nakutuma*.




```
Uwongo, Stanza 1, Line 1, H1: "ka-ra-ta-si na-ku-tu-ma"
Lexical reading: 4·4, P₁=3, P₂=7, G-lex = 2.
Auditory reading: 4·4, P₁=3, P₂=7, G-aud = 2, delta —.
Instrumental reading: 4·4, P₁=3, P₂=7, G-inst = 2, delta —
Speaker median 259 Hz; declination −0.47 st/s.
```

The syllables the instrument designates as prominent are exactly those annotated by ear: the two readings agree on the accentual structure of this line.

```
Per-syllable measurements:
  1 ka     460 ms†  F0 -2.59 st  peak +347 ms  intensity  41.1 dB
  2 ra     179 ms   F0 -2.46 st  peak + 83 ms  intensity  68.0 dB
  3 ta     240 ms   F0 +3.04 st  peak +219 ms  intensity  71.0 dB
  4 si     250 ms   F0 +2.65 st  peak +156 ms  intensity  64.8 dB
  5 na     139 ms   F0 -0.03 st  peak +  1 ms  intensity  66.4 dB
  6 ku     150 ms   F0 -1.65 st  peak + 57 ms  intensity  64.1 dB
  7 tu     189 ms   F0 +1.55 st  peak + 89 ms  intensity  71.5 dB
  8 ma     399 ms†  F0 -0.81 st  peak + 18 ms  intensity  76.1 dB
```

3. What the sample establishes

The eleven measurements above do not stand as a homogeneous block of confirmations or disconfirmations to be tallied and averaged. Three of them carry the argument of this section, each in a distinct way, and each must be read against the specific per-syllable data that produced it rather than against a general expectation of what an instrument should do. The reasoning proceeds in three stages: first, a single hemistich that destroys the bipartite axiom outright, independent of any question about instrumentation; second, a

hemistich that exposes the physical instability of the one position the accentual school treats as secure; third, a hemistich that shows, mechanism intact, how an instrument can manufacture the school's own ideal configuration out of nothing but the physics of a short syllable.

3.a. The triple convergence: an axiom destroyed without appeal to instrumentation

Figure 3, *Uwongo*, stanza 1, line 3, second hemistich, *kulla jambo kwa ndiaye*, returns the same reading on every one of the three independent procedures this study employs. The lexicon, derived from the printed text by the penultimate rule alone, gives 2·2·4, three accentual groups at positions 1, 3 and 7. The ear, screening the formal reading without reference to the lexicon, returns the identical structure: 2·2·4, positions 1, 3, 7, no nucleus marked weak. The instrument, applying no linguistic knowledge whatever and reporting only the physical convergence of F0, intensity and duration, returns 2·2·4 a third time, at the same three positions, with no divergence recorded against either of the prior readings.

This is not a case to be weighed against its exceptions. A structure independently confirmed by a deterministic rule applied to the orthography, by trained audition applied to a recording, and by a blind acoustic algorithm applied to the same recording, is not a probabilistic tendency awaiting further data. It is a fact about this hemistich, established three times over by three procedures that share no assumptions and could not have colluded, and the fact is that an eight-syllable segment of classical Swahili verse carries three accentual nuclei, not two. The bipartite axiom asserts that G = 2 is the segment's generative condition. This hemistich is not an exception the axiom must absorb; it is a direct instance of the configuration the axiom declares impossible, verified beyond the reach of the annotator's ear alone. Selected precisely because it carries the highest G attested in a corpus otherwise the most favorable to the accentual model in this entire study, its confirmation at every level available to this protocol closes, rather than opens, the question of whether G = 2 is a law. The same result has been obtained three times elsewhere in this chapter, always on the same configuration, 2·2·4 at positions 1, 3 and 7: in section 9.3, on the second hemistich of *Kiswahili!!* 3.1, *wala sina punguwani*, and in both *utenzi* samples, at *Rasi 'lGhuli* 12.4, *hiyo siku ya nushuri*, and at *Utendi wa Haudaji* 37.3, *kuwa mimi muombezi*. The configuration the axiom forbids is thus confirmed physically four times, in four poems of two meters.

3.b. The terminal anchor exposed: *Kiswahili!!* 2.4, H1

The accentual school's confidence in the eight-syllable line rests on more than the *Grenzschema*'s ideal volumes; it rests on an assumption about where the final accent of each eight-syllable segment must fall. Because the classical *shairi* line prohibits enjambment, and because no word crosses its caesura, a word boundary always coincides with the end of each hemistich, and the penultimate rule accordingly guarantees a lexical stress on the seventh syllable of each. The accentual school treats this as a metrical beat, the second post of its bipartite structure. What it is, in fact, is a passive phonological reflex of a rule about word boundaries, not a generative claim about rhythm — and the raw signal, asked to locate this supposedly secure anchor without the correction a linguist's training applies, does not reliably find it where the rule predicts.

Figure 2, *Kiswahili!!*, stanza 2, line 4, first hemistich, *kosa langu kosa gani*, is the sharpest instance of this instability available in the sample. The lexicon yields 2·2·2·2, four accentual groups at positions 1, 3, 5 and 7 — every word of the line is disyllabic, and the penultimate rule places a stress on the first syllable of each. The ear confirms the same four positions, marking the first and third as weak but not removing them from the count: G-aud = 4, unchanged from the lexicon. The instrument returns 4·4, with peaks at position 3 and position 8. The first of these, position 3, coincides with a position the lexicon already predicts. The second does not: position 8 is the hemistich's final syllable, at the caesura, one syllable beyond the penultimate position 7 that both the lexicon and the ear identify as its terminal stress.

The per-syllable measurements make the mechanism explicit. The seventh syllable, *ga*, registers a detrended F0 of +0.25 semitones, with its own measured peak arriving 303 milliseconds after onset — the longest latency recorded anywhere in this sample, itself evidence that whatever pitch movement is occurring on this syllable is not reaching its local maximum within the syllable's own span. The eighth and final syllable, *ni*, registers +1.29 semitones, higher than *ga*'s own value, and this is the syllable the instrument accordingly selects. What the accentual school calls the terminal beat, guaranteed by the word boundary, is here a rising gesture that begins on the seventh syllable and completes on the eighth — a boundary tone marking the edge of a phrase, here at the caesura and not at the end of the line, not a lexical stress on the penultimate syllable the rule predicts. The one position

the school regards as phonologically inevitable is, on the raw signal, displaced by exactly the edge effect this book has traced since section 9.3. The *utenzi* sample records the same event on a structurally identical line: at *Utendi wa Haudaji* 32.4, *kula yambo moya moya*, four groups of two syllables on lexicon and ear, the instrument again moves the final peak from the seventh syllable to the eighth (section 9.4.2).

3.c. The internal beat manufactured: *Kiswahili!!* 3.3, H1

If the terminal anchor is a passive reflex the school mistakes for design, the internal beat — the accentual group falling on position 1, 2, 3 or 4, depending on the composer's lexical choices at the start of the line — is the position on which the school's confirmation bias operates directly: lines whose internal stress happens to fall at position 3, yielding the *Grenzschema*'s full 4·4 volumes, are retained as confirming instances; lines whose internal stress falls elsewhere are set aside as exceptions to be explained away. Figure 8, *Kiswahili!!*, stanza 3, line 3, first hemistich, *nizee wanasiasa*, demonstrates that a signal can manufacture exactly this confirming pattern by mechanical accident, with no accentual reality behind it at all.

The lexicon gives 3·5, two groups at positions 2 and 7 — already outside the *Grenzschema*'s required volumes. The ear, screening the same reading, adds a third nucleus: 3·2·3, three groups at positions 2, 4 and 7, the middle one marked weak. Two independent, non-instrumental procedures thus agree that this hemistich carries either two or three accentual groups, and neither identifies position 3 as a stress. The instrument returns 4·4 at positions 3 and 7 exactly — the *Grenzschema* in its canonical form.

The per-syllable data show precisely how this configuration is manufactured. The second syllable, *ze*, is the syllable both the lexicon and the ear identify as stressed; it measures 140 milliseconds in duration, the shortest voiced syllable in the nucleus, and its own detrended F0 mean is a modest +0.50 semitones, with a measured peak arriving 133 milliseconds after onset — that is, at 95 percent of the syllable's own duration, its rise barely completing before the syllable ends. The third syllable, *e*, immediately following, measures 240 milliseconds and registers +2.58 semitones, by far the highest value in the nucleus, with its own peak arriving after only 73 milliseconds. The pitch gesture that *ze* initiates does not culminate within *ze*; it spills across the syllable boundary and reaches its physical maximum inside the following vowel. An algorithm reporting the

highest mean F0 within each syllabic window will accordingly select position 3, *e*, as the prominent syllable of this nucleus, not position 2, *ze* — not because *e* carries a lexical accent, which nothing in the word's morphology supports, but because the physiological consequence of the accent on *ze* is measured one syllable later than the accent itself. The weak nucleus at position 4, *wa-na*, registers a still lower value and is correctly not designated by the instrument, removing it from the auditory count of three; what remains, positions 3 and 7, reproduces the *Grenzschema*'s volumes and positions exactly, out of a peak-delay artifact on a stressed syllable too short to contain its own physiological consequence.

This is the internal-beat fallacy exposed by its own supposed evidence. A school that isolates lines yielding $P_1 = 3$ as confirmation of its prototype, while treating lines yielding $P_1 = 1$, 2 or 4 as exceptions, has no means of distinguishing a genuine instance of its target from an instance manufactured by exactly this mechanism — a short stressed syllable whose pitch rise completes after the syllable has ended. The confirming pattern the school would retain from this line is, on the evidence of the lexicon and the ear alike, not there at all.

3.d. The consolidated pattern

The eight hemistiches where the instrumental reading departs from the auditory one show the same few kinds of departure recorded in the other samples of this chapter, and several show more than one. In three, the final peak moves from the seventh syllable to the last: at *mbona mwanipija zitha* at the end of the line, and at *kosa langu kosa gani* and *warehemu wana*, both first hemistiches, at the caesura, which is consistent with the reader closing each hemistich as a phrase. In five, a prominence moves by one syllable. Where it moves later, the ear's syllable is followed by an identical vowel in hiatus, as in *nizee* and *kunyamaa*, so that the pitch movement completes on the second vowel, or the two values are a near-tie, three hundredths of a semitone at *mbona*; where it moves earlier, twice in *Waji Waji*, the syllable the ear chose is the longer of the two. The remaining departures add or drop an initial prominence, or drop a nucleus the ear had marked as weak. None of these patterns is specific to the *shairi*, and none involves a property of the meter.

3.e. What this establishes

The syllabic model is untouched by any of these eleven measurements: the count of syllables and the architecture of the rhyme are properties of the printed text, established before any recording is made and confirmed independently of it. The accentual model is tested here on ground it never occupied historically, and the test yields two results of unequal weight. The first is that its foundational axiom is not merely statistically weak but directly falsified, on a hemistich where deterministic derivation, trained audition and blind acoustic measurement converge without qualification on a structure the axiom declares impossible. The second is that the two components of its idealized target are independently unstable under the same physical scrutiny: the terminal position it treats as phonologically guaranteed is displaced by a boundary tone in the sample's sharpest instance, and the internal position it treats as diagnostic of a generative prototype is reproduced, volumes and positions both, by a peak-delay artifact on a line two independent human judgments agree does not carry it. A target assembled from one component that the signal will not reliably confirm and another that the signal can manufacture without cause, tested against a line the same signal independently and unambiguously contradicts, has not survived this sample. It has been shown, at each of its three critical points, not to correspond to anything the data require.

9.6.5. Conclusion on Shairi Meter

Three corpora have been examined in this meter, chosen to bracket the space of what a *shairi* can be rather than to converge on a single convenient case. The *Waji Waji* offers a fifteen-syllable *takhmisa*, 6 + 9, arithmetically closed to the tetrasyllabic template. *Kiswahili!!* offers the standard sixteen-syllable *tarbia*, 8 + 8, read by its own author, and a stricter form of *kupindua kina kimoja* than the meter's own theorists had described. *Uwongo* offers the same 8 + 8 frame under the additional discipline of *Mashairi ya Vidato*, composed by the figure the tradition names as its master. Sixty-four, seventy-two, and sixteen lines were mapped in the three corpora respectively; three hundred and one hemistiches were retained across them, after exclusions declared and justified line by line in each case study. The figures below are those of the three case studies, restated together at a finer grain than their own conclusions give individually, and the diversity of architecture is treated as part of the finding rather than a complication to be averaged away.

1. The syllabic model

Condition 1. The syllabic constant holds in 57 of 64 lines in the *Waji Waji*, 89.1 percent, and in 139 of 141 hemistiches in *Kiswahili!!*, 98.6 percent; it holds without exception in all thirty-two hemistiches of *Uwongo*. Every departure in the first two corpora is a declared *guni* of the source text, verified against the printed edition and retained in the pool rather than discarded — six of excess or deficiency and one of deficiency in the *Waji Waji*, two further instances of seven-syllable hemistiches in *Kiswahili!!*. None arises from the execution.

Condition 2. In the *Waji Waji*, two local rhyme chains and one global constant, *-ma*, hold in their strict syllabic form across all fourteen stanzas; the matrix additionally requires a permutation of the local rhyme to the caesura of the fifth line, which succeeds in fourteen stanzas of fourteen. In *Kiswahili!!*, the architecture proves stricter than the meter's generic matrix anticipates: what a relaxed reading would treat as a stanza-local variable is here a second global constant, *-ni*, locked to both the terminus of three lines and the caesura of the fourth, holding across all eighteen stanzas without exception, alongside a further constant, *-ta*, closing every fourth line. In *Uwongo*, four of five rhyme chains hold in strict syllabic form; the fifth is contradicted once, at 3.3, where the vocalic license is tested in turn and holds. The global constant, *-ngo*, closes every fourth line of every stanza without exception.

Condition 3. No word crosses a caesura in any of the three hundred and one retained hemistiches, across all three corpora.

Condition 4. Three distinct chaining mechanisms are tested, and each is confirmed without exception on every transition available to it: the simple rhyme permutation of the *Waji Waji*, verified on fourteen stanzas; the doubled global constant of *Kiswahili!!*, verified on eighteen; and the recursive rhyme propagation of *Uwongo*'s *Mashairi ya Vidato*, confirmed term by term against its own summative equation on all three of its stanza transitions. *Uwongo* additionally satisfies, at each of those same transitions, a constraint the algebraic notation does not itself carry: the verbatim repetition of a whole hemistich across the stanza boundary, confirmed by direct textual comparison rather than by any accentual measure.

Verdict. The syllabic matrices describe these three corpora exhaustively, across three different stanzaic architectures and three different rhyme mechanisms, with every exception named and none silently absorbed.

2. The accentual model

2.a. Is the bipartite axiom verified?

The *Waji Waji* supplies no eight-syllable segment in a regular line, and the axiom is accordingly tested by hemistich rather than consolidated, since consolidation conceals what the data show. In the six-syllable hemistich, the lexicon yields G = 2 in 56 of 64 lines, 87.5 percent — Sacleux's own account verified with precision — with G = 1 in six and G = 3 in two; the reading realizes the identical distribution. In the nine-syllable hemistich, the lexicon yields G = 3 in 46 of 64, 71.9 percent, with G = 2 in nine and G = 4 in nine; the reading yields G = 3 in 44, G = 2 in fourteen, G = 4 in six. The same speaker, in the same line, produces two accentual groups in the first half and three in the second, sixty-four times over: the number of groups tracks the length of the container, not a constant the meter imposes. In *Kiswahili!!*, G = 2 holds in 79 of 141 hemistiches lexically, 56.0 percent, and 86, 61.0 percent, auditorily, against G = 3 in 60 and 54 respectively and G = 4 in two and one. In *Uwongo*, G = 2 holds in 24 of 32, 75.0 percent, lexically and 23, 71.9 percent, auditorily — the highest rate recorded in this study — against G = 3 in eight and nine.

The transition. The reading departs from the lexical structure in 10 of 128 hemistiches in the *Waji Waji*, 7.8 percent; in 9 of 141 in *Kiswahili!!*, 6.4 percent; and in 1 of 32 in *Uwongo*, 3.1 percent — all three below the 12.5 percent recorded for the prose control corpus of section 9.7.

Verdict. The axiom is not verified in any corpus of this meter. It is contradicted in no fewer than one hemistich in four throughout, including in *Uwongo*, the corpus most favorable to it. In the *Waji Waji*, the same line produces a different G in each of its two halves, according to the length of that half alone — a result a fixed generative count cannot accommodate without ceasing to be fixed.

2.b. Is the *Grenzschema* a cognitive prototype?

The sequence 4·4 does not occur once in the *Waji Waji* on the lexical measure and occurs three times on the auditory one, and each occurrence is traceable to a specific defect rather than to composition: at 1.4, a second hemistich of eight syllables where the matrix prescribes nine, itself a declared *guni* of the text; at 3.1, an extra syllable articulated in *rehemayo* that breaks the count; at 11.3, a false start compounded by a distortion of the text. All three lines are

excluded from the pool as *guni* of reading, and the configuration accordingly does not occur once in a regular line of this poem. In *Kiswahili!!*, it occurs in 29 hemistiches lexically, 20.6 percent, and 34 auditorily, 24.1 percent; among the hemistiches satisfying the bipartite axiom, the extended parity criterion holds in 46.8 percent lexically and 52.3 percent auditorily, a near-even split. In *Uwongo*, it reaches its highest recorded rate in this study, 34.4 percent lexically and 31.2 percent auditorily — a plurality among the hemistiches satisfying the bipartite axiom, 45.8 percent of them, not a majority — and the extended parity criterion divides at 50.0 percent lexically and 47.8 percent auditorily, as close to chance as a finite corpus permits.

Verdict. The *Grenzschema* is not a cognitive prototype in any corpus of this meter. It is wholly absent from every regular line of the corpus least favorable to it, and reaches barely one hemistich in three even in the corpus most favorable to it — a ceiling, not a threshold crossed, and in the *Waji Waji*'s case a total exclusion traceable in every instance to a named defect of text or of reading, never to composition.

2.c. Is there a rhythm?

Nuclei of equal volume occur in 24 of 128 hemistiches in the *Waji Waji* lexically, 18.8 percent, and 23 auditorily; in 31 of 141 in *Kiswahili!!*, 22.0 percent; and in 11 of 32 in *Uwongo*, 34.4 percent — the ceiling recorded in this study, and still a minority. Across consecutive hemistiches, the recurrence of identical sequences falls at or below the rate independent draws from each corpus's own distribution would produce, in every corpus where the two hemistich positions being compared are of equal length: in *Kiswahili!!*, 11.4 percent lexically against 12.7 percent expected, and 12.9 percent auditorily against 13.7 percent expected; in *Uwongo*, 16.1 percent on both measures against 21.3 and 19.5 percent expected respectively. The *Waji Waji*'s two hemistich positions differ in length within the line and cannot be compared to each other on this measure; no recurrence beyond the six-syllable and nine-syllable sequences already given above is available to report.

Verdict. There is no rhythm at either scale, in any corpus of this meter. Equal-volume nuclei remain a minority throughout, reaching one hemistich in three at most, and where hemistich positions are comparable in length, their sequence recurs at or below the rate independence would produce.

2.d. Is the tetrasyllabic template used?

The three corpora sit at three different points on this condition by construction. In the *Waji Waji*, the template has almost no object: three hemistiches of 128, 2.3 percent, have a length divisible by four, and all three are *guni*; a tetrasyllabic nucleus of some kind occurs in 82 hemistiches, 64.1 percent, but in none does it exhaust the segment, since the arithmetic of a six- or nine-syllable container forbids it. In *Kiswahili!!*, a tetrasyllabic nucleus occurs in 54 of 141, 38.3 percent, and only 29 are partitioned entirely into two such units. In *Uwongo*, the template is arithmetically available throughout, both hemistiches measuring eight syllables in every line.

Verdict. Where the template has no arithmetical object, as in the *Waji Waji*, it produces no prediction at all for 97.7 percent of the poem's hemistiches. Where it is fully available, as in *Uwongo*, the condition is undecided by arithmetic alone, and the burden falls entirely on the three questions above, which it does not survive.

3. Two independent controls

Each corpus carries a check on the annotation itself beyond the lexical-auditory transition already reported. In *Kiswahili!!*, two annotation passes made under different notations at different times coincide on 129 of 141 comparable hemistiches, 91.5 percent; of the twelve divergent hemistiches, G falls between passes in five, rises in two, and remains equal with a changed position in five. In *Uwongo*, the two passes coincide on 30 of 32, 93.8 percent, the higher of the two rates recorded in this meter; of the two divergent hemistiches, one records an added accentual group in the second pass and one records the same count at a changed position. A further, independent check exists for *Kiswahili!!* alone: a complete IPA transcription of the recitation, made by ear with a second linguist prior to and independently of this study's own annotation, and deposited alongside the recording rather than tabulated here.

4. The instrumental sample

Eleven hemistiches were measured across the three corpora, six chosen for maximal contradiction of the model, two drawn at random, and three for the *Grenzschema* realized without qualification. The full analysis is reported in section 9.6.4, and only its bearing on this conclusion is restated here.

The instrument is offered strictly for information, applied to a small sample, and never generative: no version of the accentual model ever

employed one, and no meter is extracted from a signal in this study or anywhere in it. Within that limit, the sample yields one result that requires no instrument to already stand — a triple convergence of the lexicon, the ear and the signal on G = 3 in *Uwongo*, the corpus most favorable to the model, on the very configuration the bipartite axiom declares impossible — and two results that expose the instrument's own limits when read without the correction a trained ear applies: a terminal position the school treats as phonologically secure, displaced by a boundary tone; and an internal position it treats as diagnostic of its prototype, manufactured by a peak-delay artifact on a line two independent human judgments agree does not carry it.

5. What the shairi *meter establishes*

The three corpora examined here are not variations on a single case. They differ in stanza length, in hemistich length, in rhyme mechanism, and in the identity and status of their composers, from a nineteenth-century Islamic scholar composing a prayer for rain, to a twentieth-century author reading his own composition, to the figure the tradition names as its master. The syllabic model accounts for all three without adjustment beyond the *guni* the tradition itself defines. The accentual model's rate of apparent success rises and falls across the same three corpora, and what it tracks, on the evidence of this chapter, is not architecture but lexicon: a poet whose diction happens to favor words of two and four syllables will produce more configurations resembling the model's target, in exactly the corpus and exactly the proportion his vocabulary allows, independent of any generative rule. Where the same poem supplies two hemistich lengths at once, as the *Waji Waji* does, the number of accentual groups changes with the length of the segment within the same line, by the same speaker, which a fixed generative count cannot accommodate without ceasing to be fixed.

A model tested this exhaustively, on its least favorable ground and on its most favorable, and found each time to account for a minority of what it predicts, has been given every hearing this study could construct. It has not survived one of them.

Two claims specific to this meter can accordingly be closed here. Does the *shairi* line rest on a bipartite accentual axiom, G = 2, as a generative law? No: the axiom is contradicted in no fewer than one hemistich in four in every corpus examined, including the one most favorable to it, and in two hemistiches, from two different poems,

Uwongo and *Kiswahili!!*, the contradiction is confirmed by three independent procedures. Does Vierke's *Grenzschema* function, as claimed, as a cognitive prototype underlying the meter, or Minerba's trochaic innovation as the rhythmic principle it is said to be? No: the configuration reaches its highest recorded rate at one hemistich in three, in the corpus best suited to produce it by the sheer density of tetrasyllabic words in its diction, and it is absent from every regular line of a classical corpus this study has no reason to consider marginal. A prototype that a plurality of hemistiches contradicts, in the meter's own most favorable exemplar, is not generating that meter.

9.7. The Prose Control Corpus

9.7.1. Why This Control Exists

Conditions 5 and 6 of section 9.2.4 ask whether a given accentual configuration occurs in classical verse more often than chance would produce. A claim of that form has no measurable content until *chance* is defined, and it cannot be defined in the abstract: what has to be established is not an arbitrary probability but what the Swahili lexicon actually delivers when it is made to fill an eight-syllable container for reasons that have nothing to do with meter.

This is what the corpus below supplies. If the distribution of accentual configurations in the *utenzi* proves indistinguishable from the distribution obtained in ordinary prose confined to the same syllabic frame, then the configurations the accentual school identifies as metrical are properties of Swahili word structure inside eight syllables, and would be found in a shopping list. If the poetic distribution departs from this baseline, the accentual school has a genuine phenomenon to account for, and this study will say so.

The comparison also bears on the auditory measure. The corpus was read aloud by the same speaker (Fatuma Shaaban), under the same instruction not to sing, so that the reduction of G-lex to G-aud observed in the poem can be set beside the reduction observed in prose. Should a reader realize two accentual groups in prose as often as in verse, the bipartite realization is a fact about Swahili intonation and not about the *utenzi*.

9.7.2. How the Corpus Was Built

Language and register. The eighty lines are in Standard Swahili (*Kiswahili sanifu*) in its contemporary form, composed for this study. The register is deliberately ordinary: statements about everyday matters, short nominal lists, simple determinative phrases. No Kiamu form appears, no archaism, no elevated Arabic-derived vocabulary of the kind the classical corpus carries. Verbal forms are varied across persons and tenses so that no morphological template recurs.

Syllabic constraint. Every line contains exactly eight *mizani*, verified syllable by syllable. This constraint is artificial and is the one feature the corpus shares with the poem; without it no comparison would be possible, since the accentual configurations available in a line depend on its length.

Accentual design. The distribution of G-lex was fixed by construction rather than left to chance. There exist exactly thirteen partitions of an eight-syllable line into accentual nuclei of at least two syllables, and all thirteen are represented. G-lex = 3 and G-lex = 4 are deliberately over-represented relative to their frequency in the poem, because they carry the test: the question is whether a prose reading reduces such lines to two prominences. The *Grenzschema* sequence 4·4 is represented by eight lines, enough to show whether it behaves differently from the others. It should be noted that G = 4 forces the sequence 2·2·2·2, which is the only partition of eight into four nuclei of at least two syllables.

Absence of poetic cues. No line rhymes with any other. The order was fixed by a search that admits no two consecutive lines sharing a final syllable, no two consecutive lines sharing an accentual sequence, and no three consecutive lines sharing the same G-lex. No two lines are thematically connected, so that nothing invites the reader to treat the sequence as a text.

One compromise. Ten lines use the copula *ni*, which is difficult to avoid when the sequences 2·3·3 and 3·2·3 are required without recourse to rare vocabulary. This is the only structure that recurs noticeably.

Limitations. The corpus is composed and not attested: it is not a sample of Swahili prose but a set of sentences written to a constraint. The eight-syllable frame is itself artificial, as stated above. And the reader is a literate reader familiar with Swahili poetics, which cannot be excluded as a source of recitation habit; the limitation is the same as the one declared in section 9.2.6, and it applies here with more

force, since the whole point of the corpus is that it should not be read as poetry.

Deposit. The reading sheet as supplied to the reader, without the analysis below, is deposited with the recordings in the Zenodo repository.

Designed distribution of G-lex

G = 1 in 3 lines, G = 2 in 32, G = 3 in 36, G = 4 in 9.

By accentual sequence: 8 in 3, 2·6 in 6, 3·5 in 7, 4·4 in 8, 5·3 in 7, 6·2 in 4, 2·2·4 in 6, 2·4·2 in 6, 4·2·2 in 6, 2·3·3 in 6, 3·2·3 in 6, 3·3·2 in 6, 2·2·2·2 in 9.

9.7.3. Mapping

The entries follow the conventions of section 9.4.2: the syllabified line with its syllable count, the lexical reading derived by the penultimate rule, and the auditory reading established by ear.

1 ba-ru-a-ya-ke-ni-nde-fu (8)
lexical 3·2·3 — P_1=2, P_2=4, P_3=7 — G-lex 3
auditory 3·2·3 — P_1=2, P_2=4, P_3=7 — G-aud 3 — Δ —
2 wa-me-sa-ha-u-ka-la-mu (8)
lexical 5·3 — P_1=4, P_2=7 — G-lex 2
auditory 5·3 — P_1=4, P_2=7 — G-aud 2 — Δ —
3 mwa-li-mu-wa-ngu-ni-ju-ma (8)
lexical 3·2·3 — P_1=2, P_2=4, P_3=7 — G-lex 3
auditory 3·2·3 — P_1=2, P_2=4, P_3=7 — G-aud 3 — Δ —
4 ju-zi-tu-li-ku-ta-ni-ka (8)
lexical 2·6 — P_1=1, P_2=7 — G-lex 2
auditory 2·2·4 — P_1=1, P_2=3, P_3=7 — G-aud 3 — Δ +P_2=3
5 ye-ye-a-me-le-ta-ma-ji (8)
lexical 2·4·2 — P_1=1, P_2=5, P_3=7 — G-lex 3
auditory 2·6 — P_1=1, P_2=7 — G-aud 2 — Δ −P_2=5
6 tu-li-ku-wa-tu-me-cho-ka (8)
lexical 4·4 — P_1=3, P_2=7 — G-lex 2
auditory 4·4 — P_1=3, P_2=7 — G-aud 2 — Δ —
7 sa-fa-ri-ya-si-ku-ta-tu (8)
lexical 3·3·2 — P_1=2, P_2=5, P_3=7 — G-lex 3
auditory 3·3·2 — P_1=2, P_2=5, P_3=7 — G-aud 3 — Δ —

8 tu-li-ku-ta-ni-ka-ja-na (8)
lexical 6·2 — P_1=5, P_2=7 — G-lex 2
auditory 6·2 — P_1=5, P_2=7 — G-aud 2 — Δ —
9 ka-la-mu-hi-zi-ni-za-ngu (8)
lexical 3·2·3 — P_1=2, P_2=4, P_3=7 — G-lex 3
auditory 3·2·3 — P_1=2, P_2=4, P_3=7 — G-aud 3 — Δ —
10 di-ri-sha-li-me-vu-nji-ka (8)
lexical 3·5 — P_1=2, P_2=7 — G-lex 2
auditory 3·5 — P_1=2, P_2=7 — G-aud 2 — Δ —
11 ba-ru-a-ka-la-mu-me-za (8)
lexical 3·3·2 — P_1=2, P_2=5, P_3=7 — G-lex 3
auditory 3·3·2 — P_1=2, P_2=5, P_3=7 — G-aud 3 — Δ —
12 tu-me-a-ndi-ka-ba-ru-a (8)
lexical 5·3 — P_1=4, P_2=7 — G-lex 2
auditory 5·3 — P_1=4, P_2=7 — G-aud 2 — Δ —
13 nyu-mba-ye-tu-i-me-je-ngwa (8)
lexical 2·2·4 — P_1=1, P_2=3, P_3=7 — G-lex 3
auditory 2·2·4 — P_1=1, P_2=3, P_3=7 — G-aud 3 — Δ —
14 sa-sa-wa-na-e-le-wa-na (8)
lexical 2·6 — P_1=1, P_2=7 — G-lex 2
auditory 2·6 — P_1=1, P_2=7 — G-aud 2 — Δ —
15 nyu-mba-ye-tu-i-na-mo-to (8)
lexical 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-lex 4
auditory 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-aud 4 — Δ —
16 ha-tu-ku-wa-tu-me-la-la (8)
lexical 4·4 — P_1=3, P_2=7 — G-lex 2
auditory 4·4 — P_1=3, P_2=7 — G-aud 2 — Δ —
17 ja-na-wa-li-pi-ka-cha-i (8)
lexical 2·4·2 — P_1=1, P_2=5, P_3=7 — G-lex 3
auditory 2·6 — P_1=1, P_2=7 — G-aud 2 — Δ −P_2=5
18 wa-li-po-ke-a-ba-ru-a (8)
lexical 5·3 — P_1=4, P_2=7 — G-lex 2
auditory 5·3 — P_1=4, P_2=7 — G-aud 2 — Δ —
19 wa-tu-wo-te-wa-me-cho-ka (8)
lexical 2·2·4 — P_1=1, P_2=3, P_3=7 — G-lex 3
auditory 2·2·4 — P_1=1, P_2=3, P_3=7 — G-aud 3 — Δ —
20 a-me-nu-nu-a-mi-ka-te (8)
lexical 5·3 — P_1=4, P_2=7 — G-lex 2
auditory 5·3 — P_1=4, P_2=7 — G-aud 2 — Δ —

21 tu-li-pa-nda-ba-si-ja-na (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
22 ga-ri-la-ke-li-na-ra-ngi (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
23 wo-te-wa-na-ji-ta-hi-di (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
24 ma-ma-ya-ke-a-me-fi-ka (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
25 mwa-li-mu-a-na-fu-ndi-sha (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
26 ha-ba-ri-za-le-o-hi-zi (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
27 ka-la-mu-i-me-po-te-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
28 a-na-fa-nya-ka-zi-ngu-mu (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
29 ni-li-ku-wa-ni-me-cho-ka (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
30 m-ka-te-hu-u-ni-wa-ko (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ −P1=2
31 mi-mi-ni-na-wa-e-le-wa (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
32 sa-ma-ki-m-ka-te-ma-ji (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
33 da-wa-hi-i-i-na-la-dha (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ −P3=5

34 mwa-li-mu-wa-shu-le-ye-tu (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —
35 ndu-gu-ya-ngu-a-na-ho-ma (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
36 ki-ta-bu-hi-ki-ni-cha-ke (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
37 wa-na-fu-nzi-wa-na-so-ma (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
38 ba-ba-ya-ngu-a-na-ga-ri (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
39 le-o-ba-ru-a-ni-ya-ke (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
40 wa-li-ku-wa-wa-na-i-mba (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
41 ke-sho-sa-fa-ri-ni-nde-fu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
42 a-na-ta-ka-ku-sa-fi-ri (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
43 da-da-ya-ngu-a-na-so-ma (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
44 da-da-ya-ke-a-na-ka-zi (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ −P3=5
45 wa-ge-ni-we-tu-nyu-mba-ni (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
46 ni-na-wa-e-le-wa-wo-te (8)
lexical 6·2 — P1=5, P2=7 — G-lex 2
auditory 6·2 — P1=5, P2=7 — G-aud 2 — Δ —

47 ju-zi-ha-ba-ri-ni-mba-ya (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
48 ja-na-wa-li-ku-ta-ni-ka (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —
49 wa-tu-wo-te-wa-na-nja-a (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
50 wo-te-wa-me-ka-a-chi-ni (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·4·2 — P1=1, P2=5, P3=7 — G-aud 3 — Δ —
51 ma-ma-hu-yu-a-na-pe-sa (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
52 wa-li-ku-wa-ha-pa-ja-na (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
53 a-me-fu-ngu-a-di-ri-sha (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
54 ka-ka-ya-ke-a-na-che-za (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
55 m-ge-ni-a-me-o-ndo-ka (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
56 sa-sa-sa-ma-ki-ni-gha-li (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4?, P3=7 — G-aud 3 — Δ —
57 ma-ga-ri-ya-me-si-ma-ma (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
58 tu-me-we-ka-pe-sa-be-nki (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
59 m-vu-la-na-a-me-la-la (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —

60 ha-tu-ku-wa-pe-le-ke-a (8)
lexical 8 — P1=7 — G-lex 1
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ +P1=4
61 ni-me-o-na-ga-ri-la-ke (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
62 ha-wa-ku-si-ki-a-vye-ma (8)
lexical 6·2 — P1=5, P2=7 — G-lex 2
auditory 6·2 — P1=5, P2=7 — G-aud 2 — Δ —
63 ha-wa-ku-wa-e-le-we-sha (8)
lexical 8 — P1=7 — G-lex 1
auditory 8 — P1=7 — G-aud 1 — Δ —
64 ha-ba-ri-zi-me-e-ne-a (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
65 ke-sho-tu-ta-so-ma-sa-na (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·4·2 — P1=1, P2=5, P3=7 — G-aud 3 — Δ —
66 a-li-nu-nu-a-m-ka-te (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
67 ba-ba-ya-ngu-a-me-la-la (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
68 m-na-ta-ka-ku-o-ndo-ka (8)
lexical 4·4 — P1=3, P2=7 — G-lex 2
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ —
69 ha-pa-m-vu-a-ni-nyi-ngi (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
70 wa-me-ji-ta-hi-di-sa-na (8)
lexical 6·2 — P1=5, P2=7 — G-lex 2
auditory 8 — P1=7 — G-aud 1 — Δ −P1=5
71 ha-pa-cha-ku-la-ni-che-tu (8)
lexical 2·3·3 — P1=1, P2=4, P3=7 — G-lex 3
auditory 2·3·3 — P1=1, P2=4, P3=7 — G-aud 3 — Δ —
72 le-o-wa-me-ji-ta-hi-di (8)
lexical 2·6 — P1=1, P2=7 — G-lex 2
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ —

73 ha-pa-tu-na-pi-ka-cha-i (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ −P2=5
74 tu-na-ta-fu-ta-sa-ma-ki (8)
lexical 5·3 — P1=4, P2=7 — G-lex 2
auditory 5·3 — P1=4, P2=7 — G-aud 2 — Δ —
75 na-ha-wa-ku-tu-e-le-wa (8)
lexical 8 — P1=7 — G-lex 1
auditory 3·5 — P1=2?, P2=7 — G-aud 2 — Δ +P1=2
76 le-o-tu-me-a-nza-ka-zi (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·4·2 — P1=1, P2=5, P3=7 — G-aud 3 — Δ —
77 wa-ge-ni-wa-me-wa-si-li (8)
lexical 3·5 — P1=2, P2=7 — G-lex 2
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ —
78 wa-me-ka-a-ha-pa-sa-na (8)
lexical 4·2·2 — P1=3, P2=5, P3=7 — G-lex 3
auditory 4·2·2 — P1=3, P2=5, P3=7 — G-aud 3 — Δ —
79 ka-ka-ya-ngu-a-na-du-ka (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
80 di-ri-sha-la-nyu-mba-ye-tu (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ —

9.7.4. Results

The figures below describe the prose corpus and nothing else. They are the baseline; the comparisons that use it are made in the case studies, each against its own data.

Distribution of G. G-lex = 1 in 3 lines, G-lex = 2 in 32, G-lex = 3 in 36, G-lex = 4 in 9. After the reading, G-aud = 1 in 2 lines, G-aud = 2 in 36, G-aud = 3 in 35, G-aud = 4 in 7.

These two distributions are reported for completeness and are not themselves comparable with any corpus. The distribution of G-lex was chosen when the corpus was written, as section 9.7.2 states, and it is a property of the design. What follows is not.

Transition from G-lex to G-aud. The reading departs from the lexical structure in **10 of the 80 lines, that is 12.5 percent**. There are 7 unrealized nuclei and 3 added prominences. Of the 45 lines

whose lexicon yields more than two accentual groups, 4 are reduced to two, that is 8.9 percent. Of the 32 lines whose lexicon already yields two, 30 are realized with two.

This is the quantity the baseline exists to establish. Reading ordinary Swahili prose aloud, a native speaker alters the accentual structure supplied by the lexicon in roughly one line in eight, under no metrical constraint whatever and with no poetic text before her. Any corpus whose transition rate approaches this figure is exhibiting what Swahili reading does to an eight-syllable string, and not what a meter does to it.

The sequence 4·4 in prose. Eight lines have the lexical sequence 4·4, and all eight are realized with $P_1 = 3$ and $P_2 = 7$. The configuration identified by the accentual school as the cognitive prototype of the *utenzi* therefore arises automatically in ordinary prose whenever the lexicon supplies two words of four syllables. It requires no metrical intention, and in this corpus it never fails to appear once the lexical conditions for it are met.

The converse also holds here: no line whose lexicon yields a different sequence is realized as $P_1 = 3$, $P_2 = 7$. In this corpus the configuration appears when the words produce it and never otherwise.

Positions of the two prominences. Among the 36 lines realized with two accentual groups, the prominences fall on positions 4 and 7 in 9 lines, on 1 and 7 in 8, on 3 and 7 in 8, on 2 and 7 in 8, and on 5 and 7 in 3. The second prominence is invariably on position 7, which follows from the penultimate rule applied to the final word; the first is dispersed across positions 1 to 5, and no position among them is favored.

Use of this baseline. Each case study reports its own transition rate, its own frequency of the sequence 4·4 and of its realization, and its own distribution of first-prominence positions, and sets them beside the figures above. The distributions of G-lex are never compared, for the reason given in section 9.7.2; only the transition, the realization of 4·4, and the dispersion of positions are.

9.8. The Marked-Syntax Control

9.8.1. Why This Second Control Exists

The prose corpus of section 9.7 is written in canonical word order. Classical Swahili verse is not: it displaces objects, adjectives and adverbs before the verb with a freedom that ordinary prose does not permit. An objection follows, and it must be met. If the transition from G-lex to G-aud is measured in canonical prose and compared with verse, the similarity or difference observed might be an effect of word order rather than of meter.

Twenty further sentences were therefore composed, in the same everyday vocabulary and under the same eight-syllable constraint, but with the marked constituent orders characteristic of the classical corpus: object before verb, adjective before noun, adverb in initial position. Three of them are single verbal words of eight syllables, which the *utenzi* also produces. They were read by the same speaker (Fatuma Shaaban), under the same instruction not to sing.

These twenty sentences are **not** part of the null baseline and are never added to it. Their function is narrower: to establish whether the transition rate measured in section 9.7 depends on word order. They are compared with the baseline, not merged into it.

9.8.2. Mapping

1 nya-ma yo-te hu-la pa-ka (8)
lexical 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-lex 4
auditory 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-aud 4 — Δ — *major on P_2 and P_4, minor on P_1 and P_3*
2 we-ma sa-na wa-tu ha-wa (8)
lexical 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-lex 4
auditory 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-aud 4 — Δ — *major on P_2 and P_4, minor on P_1 and P_3*
3 tu-nda hi-li ye-ye hu-la (8)
lexical 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-lex 4
auditory 2·2·2·2 — P_1=1, P_2=3, P_3=5, P_4=7 — G-aud 4 — Δ — *no audible major/minor contrast*

4 ma-ji ha-ya ku-nywa sa-sa (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ — *major on P2 and P4, minor on P1 and P3*
5 ka-zi zo-te fa-nya we-we (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ — *major on P2 and P4, minor on P1 and P3*
6 pe-sa ze-tu wa-tu hu-la (8)
lexical 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-lex 4
auditory 2·2·2·2 — P1=1, P2=3, P3=5, P4=7 — G-aud 4 — Δ —
7 cha-ku-la hi-ki ka-le-ta (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
8 ki-ta-bu cha-ko na-so-ma (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
9 ma-ta-mu sa-na ma-tu-nda (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ —
10 ki-su hi-ki ni-me-twa-a (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
11 ga-ri le-tu wa-me-i-ba (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
12 ha-pa wa-me-fi-ka wa-tu (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·4·2 — P1=1, P2=5, P3=7 — G-aud 3 — Δ —
13 ka-zi wa-na-fa-nya sa-sa (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ –P2=5?
14 du-ka le-tu a-me-fu-nga (8)
lexical 2·2·4 — P1=1, P2=3, P3=7 — G-lex 3
auditory 2·2·4 — P1=1, P2=3, P3=7 — G-aud 3 — Δ —
15 sa-na wa-na-pe-nda nya-ma (8)
lexical 2·4·2 — P1=1, P2=5, P3=7 — G-lex 3
auditory 2·6 — P1=1, P2=7 — G-aud 2 — Δ –P2=5?

16 wa-zu-ri wa-ge-ni ha-wa (8)
lexical 3·3·2 — P1=2, P2=5, P3=7 — G-lex 3
auditory 3·3·2 — P1=2, P2=5, P3=7 — G-aud 3 — Δ — *P2 may be a minor stress*
17 nyu-mba-ni kwe-tu ka-la-la (8)
lexical 3·2·3 — P1=2, P2=4, P3=7 — G-lex 3
auditory 3·2·3 — P1=2, P2=4, P3=7 — G-aud 3 — Δ — *three major stresses, no minor*
18 ha-tu-ta-wa-ka-ri-bi-sha (8)
lexical 8 — P1=7 — G-lex 1
auditory 4·4 — P1=3, P2=7 — G-aud 2 — Δ +P1=3 *the added nucleus falls on the tense marker -ta-*
19 a-li-po-tu-fa-ha-mi-sha (8)
lexical 8 — P1=7 — G-lex 1
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ +P1=2 *the added nucleus falls on the tense marker -li-*
20 wa-li-yo-tu-fu-ndi-shi-a (8)
lexical 8 — P1=7 — G-lex 1
auditory 3·5 — P1=2, P2=7 — G-aud 2 — Δ +P1=2 *the added nucleus falls on the tense marker -li-; a pause follows waliyo*

9.8.3. Results

Transition. The reading departs from the lexical structure in 6 of the 20 sentences, that is 30 percent. Three of those six are the single-word sentences, in which the lexicon yields one nucleus only, so that any second prominence is an addition by definition. Excluding them, the rate is 3 of 17, that is 17.6 percent, against 12.5 percent in the canonical corpus of section 9.7.

Marked word order therefore raises the departure rate somewhat, on numbers too small to support a statistical claim. What it does not do is change the order of magnitude: a reader still realizes the lexical structure of most lines exactly, whether the constituents stand in canonical order or not. The comparison of section 9.7 with any poetic corpus is not vitiated by word order.

The single-word sentences, and where the added nucleus falls. In all three cases the added prominence lands on the tense marker: on *-ta-* in *hatutawakaribisha*, on *-li-* in *alipotufahamisha* and in *waliyotufundishia*. The position is not arbitrary and it is not metrical: it is the morphological slot that carries tense, and the reader

marks it. In the last of the three, a pause follows *waliyo*, and the waveform shows the sentence divided into two distinct stretches of sound separated by a silence, the first ending on the relative marker.

One of these produces the *Grenzschema*. In *hatutawakaribisha* the added prominence on *-ta-* falls on the third syllable, and the lexical stress on the seventh. The result is P1 = 3, P2 = 7, with two accentual groups: the configuration the accentual school identifies as the cognitive prototype of the *utenzi*. It arises here in a single prose word, from the placement of a tense marker, with no line, no rhyme and no meter of any kind.

A graded hierarchy of prominence. The most consequential observation of this control concerns the six sentences whose lexicon yields four accentual groups in the sequence 2·2·2·2. In four of them the four prominences are not equal: those on positions 3 and 7 are clearly stronger, and those on positions 1 and 5 are weaker. In one they are indistinguishable. In others across the corpus a nucleus can be heard as major or minor according to its contrast with its neighbors, and the assignment varies from sentence to sentence.

Two consequences follow, and they point in the same direction.

The first is that this is precisely the structure Harries described in 1962, when he recorded that Swahili verse contains "major (usually penultimate) and minor stresses" while denying that they constitute a metrical system of stress. It is also, position for position, the four-beat model Jones proposed in 1976, with major beats on the third and seventh syllables and minor beats on the first and fifth. Both descriptions are confirmed here — in prose, in sentences about food and visitors, composed for this study and never before uttered.

The second is that a two-beat reading of such a line is available at any moment. Attending only to the major prominences yields two beats on positions 3 and 7; attending to all four yields the full lexical structure. The reader's own report is that two major stresses are generally clear, that minor ones are audible or not according to contrast, and that the assignment is variable. An analyst listening for two beats will find two, and one listening for four will find four, in the same utterance. That is not a property of the *utenzi*: it is a property of how Swahili prose is heard.

9.9. Conclusion: What the Three Meters Establish

Eight case studies have been conducted across three meters — the *utenzi*, the *wimbo*, and the *shairi* — under the same protocol throughout: a lexical reading derived deterministically from the printed text by the penultimate rule, an auditory reading screened by ear from a formal recitation, and, on a reasoned and randomly supplemented sample in each meter, an instrumental reading offered strictly for information and never as a means of extracting a meter from a signal. Two hundred and ninety lines were retained in the *utenzi*, twenty-eight lines comprising sixty hemistiches in the *wimbo*, and a hundred and fifty-two lines comprising three hundred and one hemistiches in the *shairi* — four hundred and seventy lines in total. The detailed findings of each meter are stated in their own conclusions, 9.4.3, 9.5.5, and 9.6.5, and are not repeated here in full. What follows draws the three together, condition by condition, and states what the chapter as a whole has established, with the precision the data allow and no further.

9.9.1. The Syllabic Model, Condition by Condition

Condition 1 tests whether the absolute syllable count of each line or hemistich matches the number the formal matrix prescribes. It holds in every corpus of every meter beyond a small number of departures, and every one of those departures is named, verified against the printed source, and classed as *guni* — a category the endogenous tradition itself defines and sanctions, not one this study introduces to protect the model. In the *utenzi*, the constant holds in 282 of 290 lines, 97.2 percent, the eight exceptions distributed across both corpora and confirmed against the printed editions. In the *wimbo*, it holds without a single exception across all sixty hemistiches of the three corpora tested, including the hyper-fractionated matrix, where a fifteen-syllable line carries three internal boundaries and none is breached. In the *shairi*, it holds in 89.1 percent of the *Waji Waji*'s lines and 98.6 percent of *Kiswahili!!*'s hemistiches, and without exception in *Uwongo*.

Condition 2 tests whether the rhyme scheme, and in particular its terminal syllable, matches the full syllabic form the matrix specifies rather than a looser match on the vowel alone. It holds in that strict

syllabic form throughout the corpus, with the looser vocalic license invoked only where a specific line requires it and named as such in every instance: once in the *utenzi*'s two poems combined, once in the *Waji Waji*, once in *Uwongo*. In *Kiswahili!!*, the rhyme architecture proves stricter than the meter's generic matrix anticipates, locking a second constant to both the terminus of three lines and the caesura of the fourth across all eighteen stanzas without exception.

Condition 3, tested wherever a meter carries an internal caesura, asks whether that boundary is a real compositional constraint or a typographic convenience a word may cross freely. It is satisfied without one word crossing one boundary, across sixty-eight boundaries in the *wimbo* and all three hundred and one hemistiches retained in the *shairi*. This condition does not apply to the *utenzi*, which carries no internal caesura.

Condition 4, tested wherever a matrix specifies a permutation of its rhyme or a chaining between stanzas, asks whether that architecture is realized exactly as the equation predicts, term by term. It is satisfied on every transition of every mechanism examined: the simple rhyme permutation of the *Waji Waji*, the doubled global constant of *Kiswahili!!*, and the recursive rhyme propagation of *Uwongo*'s *Mashairi ya Vidato*, confirmed stanza by stanza against its own summative equation on all three of its transitions. This condition does not apply to the *utenzi* or to the *wimbo* corpora tested, which employ neither permutation nor recursive chaining.

Consolidated verdict on the syllabic model. Across three meters differing in every formal respect examined — hemistich length, stanzaic architecture, rhyme mechanism, the presence or absence of an internal caesura, the presence or absence of stanzaic chaining — the model is not falsified once beyond the exceptions the tradition itself names.

9.9.2. The Accentual Model, Condition by Condition

Condition 5 asks whether an eight-syllable segment carries exactly two accentual groups, as the bipartite axiom requires. In the *utenzi*, the lexicon yields exactly two groups in 55.5 percent of lines and the reading in 60.3 percent; three groups occur in 41.0 percent of lines lexically, a second outcome of the same order of frequency as the first, not a residue. In the *wimbo*'s eight-syllable segments, the axiom holds in 54.5 percent lexically and 51.5 percent auditorily. In the *shairi*, the rate depends on which corpus supplies the eight-syllable

segment being measured: 56.0 percent lexically in *Kiswahili!!*, and 75.0 percent in *Uwongo* — the highest rate recorded in this study, and still a contradiction in one hemistich in four. The contradiction is confirmed at the level of physical measurement four times, in four poems of two meters, and always in the same configuration: at *Rasi 'lGhuli* 12.4, *Utendi wa Haudaji* 37.3, *Kiswahili!!* 3.1 and *Uwongo* 1.3, the last two in their second hemistich, the deterministic reading, the auditory annotation, and the instrumental signal converge, independently, on three accentual groups, 2·2·4, at positions 1, 3 and 7, a configuration the axiom declares cannot occur. Where the instrument departs from the two human readings on such lines, it does so the same way in both meters: at *Haudaji* 32.4 and *Kiswahili!!* 2.4, two lines built alike from four words of two syllables and carrying four groups on lexicon and ear, it moves the final peak from the seventh syllable to the eighth, at the end of the line in the first and at the caesura in the second: a boundary tone rather than an accent.

A separate control corpus bears directly on this condition. Eighty lines of ordinary Swahili prose, composed for no metrical purpose and read under the identical protocol, depart from their own lexical accentual structure at 12.5 percent — a rate the *utenzi* corpus does not exceed (11.4 percent) and several *shairi* and *wimbo* corpora fall below (3.1 to 12.1 percent). A reader modifies the accentual structure she is given at statistically the same rate whether she is reading a nineteenth-century epic or a sentence composed for this study with no poetic intention whatsoever.

Condition 6 asks whether the two groups, where they occur, fall specifically at the third and seventh syllable positions — the *Grenzschema*, proposed as the meter's ideal cognitive target. In the *utenzi*, the complete configuration — correct count and correct placement together — is satisfied in 16.9 percent of lines lexically and 19.3 percent auditorily, and its rate differs by a factor of two between the two poems examined, 11.4 percent in one against 22.7 percent in the other, a swing incompatible with a fixed generative matrix. In the *wimbo*'s eight-syllable segments, it occurs in 24.2 percent of hemistiches on both measures, level with a three-nucleus sequence the accentual model itself classifies outside the rhythm it claims to generate. In the *shairi*, the configuration is structurally absent from the *Waji Waji*, where the eight-syllable segment it requires occurs in only 2.3 percent of hemistiches, all of them *guni*; it occurs in 20.6 to 24.1 percent of *Kiswahili!!*'s hemistiches; and it reaches its highest recorded rate anywhere in this study in *Uwongo*, 31.2 to 34.4 percent

— a ceiling that still leaves two hemistiches in three without the configuration it is said to generate.

The same prose control corpus confirms that this configuration, where it appears, is not confined to verse: every prose sentence whose lexicon happens to supply two four-syllable words is realized with the identical third-and-seventh placement, and no sentence whose lexicon supplies another sequence is so realized.

Condition 7 asks whether a metrical foot can be said to exist at all, which requires both that a hemistich's accentual nuclei be of equal size and that the same configuration recur from one hemistich to the next. Neither is met in any corpus of any meter. Nuclei of equal size occur in 20.3 percent of *utenzi* lines, 27.3 percent of the *wimbo*'s eight-syllable hemistiches, and between 24.2 and 34.4 percent of *shairi* hemistiches depending on corpus — a ceiling below one hemistich in three throughout. Measured across consecutive units, the recurrence of identical configurations falls at or below what independent draws from each corpus's own distribution would produce in every instance where the test could be run: 14.6 percent against 13.0 percent expected in the *utenzi* — an excess of fewer than five pairs in two hundred and eighty-eight — and 13.8 percent against 15.6 percent expected in the *wimbo*, below chance.

Condition 8 asks whether the four-syllable foot proposed as the shared generative unit of all three meters can be constructed from the syllable lengths a given corpus actually supplies. In the *utenzi*, the template is arithmetically available in every regular line and is realized as two complete tetrasyllabic units in only 16.9 percent of them. In the *wimbo*, the template has no object at all in 36.7 percent of the corpus's sixty hemistiches — eighteen segments of two syllables, two of three, two of six — and is realized where available in 36.4 percent of the remainder. In the *shairi*, the *Waji Waji* excludes the template from 97.7 percent of its hemistiches by arithmetic alone; *Uwongo* makes the template available throughout, both hemistiches measuring eight syllables in every line, and the burden accordingly falls entirely on Conditions 5 through 7, which it does not survive.

Consolidated verdict on the accentual model. What varies across the eight case studies is not whether this model holds but by how much it falls short, and the variation itself is diagnostic rather than incidental: the rate rises where a composer's diction happens to favor words of two and four syllables, as in *Uwongo*, and falls where it does not, as in the factor-of-two difference between the *utenzi*'s own two poems. A configuration whose frequency tracks vocabulary rather

than remaining constant across compositions in the same meter is not evidence of a generative rule; it is an arithmetical consequence of lexical choice, measured here at the same rate in prose composed with no metrical intention as in the classical corpus itself.

A further point bears on how this shortfall has been received rather than on its magnitude. Minerba classifies each measured line by its adherence to a theorized norm, sorting departures into fulfilment, exception, and error (2025, p. 101), so that a line whose count departs from the expected value is recorded as an error and a structurally inconvenient hemistich as an exception carried by a prosodic contraction. A framework that sorts every departure from its own prediction into one of two discard categories cannot be contradicted by a departure, which is the condition of an unfalsifiable hypothesis rather than a tested one. What this sorting discards is, in at least one instance this study has examined directly, not marginal. The two-syllable hemistiches that open each of the *Wimbo wa Miti*'s eight stanzas are not short by default: each is closed by the stanza's own rhyme, matched against an eight-syllable hemistich elsewhere in the same stanza, without a single license, sixteen times over. The endogenous tradition does not record this configuration as a defect. Shariff introduces a comparably fractionated *wimbo* matrix specifically to show that the form can carry more rhymes than a simple count would suggest, and names its composer, Nabhany, a *bingwa* — a master — for the construction (1988, pp. 47-48). The segments that place a line beyond the reach of the imported tetrasyllabic template are, on the tradition's own testimony, not the corpus's incidental margin. They are what it calls achievement.

Taken together, the instrumental samples of this chapter — thirty-four segments across the three meters, offered for information and never as a source of the meter — show four things. The instrument agrees with the ear on the number of accentual groups in twenty-four segments and on their exact placement in eleven: it is a fair witness to how many prominences a segment carries and a poor one to where they fall. The nuclei the ear marked as weak are the ones it fails to find: one of nine is confirmed at the same position, against fifty-six of seventy-seven firm nuclei. When it reports fewer groups than the ear, as it does in eight of the ten segments where the counts differ, the groups it loses are initial or internal, never the terminal one, which at most moves to the last syllable; on the three segments carrying four groups, it keeps the third position and drops the first or the fifth, the very positions the prose controls show to be minor in ordinary speech

(section 9.8). And its single-syllable displacements run in both directions: later in five cases, four of them where the stressed syllable is a syllabic nasal or is followed by an identical vowel, so that the pitch movement completes on the next vowel, the fifth a near-tie; earlier in seven, in six of which the syllable the ear chose is the longer of the two. Pitch alone locates a Swahili stress less reliably than the ear, which weighs pitch, length and loudness together, and nothing in these figures is a property of the meter.

9.9.3. Where the Confusion Is Located

The recurring finding of this chapter is not that Swahili verse is free of configurations resembling a bipartite or symmetrical accentual pattern. It is that these configurations, wherever they occur, occur at the same rate, and by the same mechanism, in text composed with no metrical intention whatsoever, as the prose control corpus of section 9.7 establishes directly. An accent, in Swahili, is a property of a word, assigned by a fixed rule of the language's phonology and external to the two formal properties — the syllable count and the rhyme architecture — that this book's evidence shows to generate the line. A meter is a property of those two formal architectures alone. The recurring error tested across this chapter has been to treat the first as evidence for the second: a configuration produced automatically by ordinary lexical choice has been read, repeatedly, as a compositional achievement, and its absence from a given line has been classed as a defect of that line rather than as a fact about the words a composer happened to select.

The method applied throughout this chapter is next applied to a corpus assembled for a different purpose than philological transmission: a text composed, in the chapter that follows, to satisfy the accentual model by construction rather than to be found and later tested.

Chapter 10. Metapoetic Falsification: A Poem Engineered to Satisfy the Accentual Model

The preceding chapter tested the accentual model against classical texts the poets composed without reference to it, and found that model wanting on every condition it could be held to. A different test is available, and it is decisive in a different way. If a metrical theory proposes a generative law, that law should be constructible: a text composed deliberately to satisfy it should be able to do so completely. This chapter constructs such a text, verifies that it satisfies every accentual condition of this study without exception, and draws the two consequences that follow from a positive result obtained this way.

10.1. Why a Generative Claim Must Be Constructible

A metrical theory that only describes texts after the fact, absorbing whatever a corpus supplies and classifying departures as exceptions, cannot be tested by construction, because nothing a composer does could count as satisfying or failing it. A theory that proposes a generative law makes a stronger claim: that a text built to the letter of that law is possible, and that such a text represents the configuration the law is a law of. Vierke's bipartite structure, ideally locking the primary stresses of the *utenzi* line onto the third and seventh syllables, is a claim of this second kind, and she states its ideal form plainly while conceding that the classical corpus itself "hardly ever completely realises" it (Vierke, 2011, p. 37). Minerba inherits the same ideal and extends it into a diffusionist derivation covering three meters (2025, pp. 217–219).

If this ideal is a generative law rather than a retrospective label for whichever lines happen to fit it, a text can be built that satisfies it throughout, and that text can be checked against the same eight conditions applied to every corpus in this study. This chapter builds such a text and checks it.

10.2. *Utenzi wa Wasomi na Washairi* (The Poem of the Scholars and the Poets) — Mawazoni (2026)

The poem below was composed under one constraint alone: every line consists of exactly two words of exactly four syllables each. Because Swahili lexical stress falls on the penultimate syllable of a word, a four-syllable word places its stress on its own third syllable, and two such words in sequence place the line's two stresses on positions three and seven without further adjustment. No other constraint governs the composition; the constraint is sufficient by itself to force the configuration Vierke and Minerba describe as ideal.

Washairi nyamazeni	O Poets, hold your tongues
Watafiti sikizeni	Researchers, lend your ears
Wanavyuo amuruni	Academics, give the orders
Maagizo yashikeni	Hold fast to the instructions
Kuandika kupangika	To write and to organize
Na mizani kupimika	And the measures to be weighed
Karatasi kufanyika	Research papers being crafted
Elimuni jifunzeni	Within schooling, learn your lesson
Tutafunza kusongea	We'll teach them to move forward
Hawajui kutembea	For they know not how to walk
Mwelekeo kupotea	Their direction going astray
Mafundisho sikizeni	Listen closely to the teachings

Mawazoni, Paris 06/2026

The poem stages a confrontation between an academic establishment and the Swahili poets it addresses. The first stanza has researchers and academics issue instructions to silenced poets; the second has that establishment claim authority over the very apparatus of *vina na mizani* it did not devise; the third extends the claim to the poets' capacity to walk unassisted — an image chosen because it makes the imposition legible as such: an outside authority instructing a native practitioner in a competence that practitioner already possesses. The poem's own form performs the same imposition on the Swahili line that its content describes being performed on Swahili poets, and the two are meant to be read together.

10.3. Metrical Verification

The verification follows the same protocol as the eight case studies of chapter 9, applied here to a text with no recitation to screen, since the claim under test is generative rather than acoustic: what is at issue is whether the configuration can be built on the page, not how it is executed aloud. Conditions 5 through 8 are accordingly verified on the lexical reading alone, derived by the penultimate rule from the printed text exactly as in every prior chapter.

Condition 1. All twelve lines measure exactly eight syllables. No *guni* occurs.

Condition 2. The rhyme follows the standard eight-syllable *utenzi* matrix (Section 8.3; QRG.2.1). The fourth line of every stanza closes on the global constant X, *-ni*: *yashikeni*, *jifunzeni*, *sikizeni*. The first three lines of each stanza carry their own local rhyme, which changes with each stanza: *-ni* in the first, *-ka* in the second, *-a* in the third. Every rhyme of the poem is a syllabic rhyme in the strict sense of the matrix (a^s and X^s, QRG.1): the lines it binds close on the same full syllable. The poem is a well-formed *utenzi* by the criteria of chapter 8, independent of anything that follows.

Condition 3. Not applicable: the *utenzi* line carries no internal caesura.

Condition 4. Not applicable: the *utenzi* matrix prescribes no permutation of its rhyme and no recursive chaining between stanzas.

Condition 5. Every line yields exactly two accentual groups. G = 2 in 12 lines of 12, 100 percent.

Condition 6. Every line yields the sequence 4·4 with stresses at positions 3 and 7 — the *Grenzschema* in its exact canonical form. 12 lines of 12, 100 percent. One line requires the clisis rule already established in this study: at 2.2, *Na mizani kupimika*, the monosyllabic proclitic *na* attaches to *mizani*, forming a four-syllable nucleus, *na-mi-za-ni*, whose own penultimate falls at position 3; the second word, *kupimika*, supplies the remaining four syllables and the stress at position 7. Applied consistently, the rule confirms 4·4 here as everywhere else in the poem.

Condition 7. Every line's two nuclei are of equal volume, four and four. Twelve lines of twelve exhibit internal periodicity — the only corpus in this study where that figure approaches anything near universal, against a ceiling of 34.4 percent recorded in the most favorable classical corpus examined, *Uwongo* (section 9.6.3).

Condition 8. The tetrasyllabic foot is not merely available; it is realized in the entirety of the text. Twenty-four nuclei out of twenty-four are tetrasyllabic. No other corpus in this study approaches this figure; the *utenzi* corpora of chapter 9 realize it in 16.9 to 22.7 percent of lines, and the classical *shairi* and *wimbo* corpora remain lower still or arithmetically barred from it altogether.

Verdict. On every condition the accentual school proposes as generative, this poem returns the complete ideal, without a single exception, across twelve lines. What chapter 9 could not find in two hundred and ninety lines of the *Rasi 'lGhuli* and the *Haudaji*, in sixty hemistiches of three *wimbo* corpora, or in three hundred and one hemistiches of three *shairi* corpora, is fully present here, because here alone the text was built for that purpose and for no other.

10.4. Consequences for the Metrical Analysis of This Book

Two consequences follow from this result, and they support the argument of the preceding chapter in different ways.

The model is not vacuous. A theory that could never in principle be satisfied would tell us nothing by failing on real corpora, since nothing could have counted as success. This poem shows that the accentual ideal is constructible: given a lexicon restricted to four-syllable words, the *Grenzschema* is not merely available but unavoidable. The absence of this configuration from the classical corpora examined in chapter 9 is accordingly a fact about what those composers did, not an artifact of an unreachable standard. The comparison this chapter licenses is exact: nowhere in four hundred and seventy lines of philologically established nineteenth- and twentieth-century verse does the *Grenzschema* approach the rate achieved here by construction, and the gap between that rate and the historical record is the same gap chapter 9 measured corpus by corpus.

What is achieved by construction is not what the classical tradition produced. The lexical restriction sufficient to generate the ideal — every word exactly four syllables — is not a restriction any composer examined in this study ever adopted, and the reason is recoverable from the result itself. A text built under it repeats one rhythmic shape twelve times running, at the cost of the lexical range

that gives Swahili verse its resources: compounds, proclitics, words of two, three, five and six syllables, all of which chapter 9's corpora draw on freely and none of which this poem can admit without breaking its own constraint. The classical tradition's indifference to the *Grenzschema*, measured throughout chapter 9 at a ceiling of one hemistich in three even in its most favorable instance, is not a failure to reach an ideal its poets were aiming at. It is the ordinary consequence of composing in the language's full lexicon rather than in a subset selected to satisfy an imported target.

The chapter that follows draws these results, and those of chapter 9, together with the argument of the book as a whole.

Chapter 11. General Conclusion: The Dismantling of the Procrustean Bed and the Call for Open Science

A classical Swahili line is always measured one way first. It is counted, syllable by syllable, and its rhyme checked against the pattern the stanza requires — both operations anyone can repeat from the printed text alone, and reach the same result every time, whatever oral form the line's own later life takes, recited, chanted, or sung, since that text is the stable ground each of those forms departs from rather than a substitute for them. For a century and a half, a specific lineage of scholarship has not denied this count; it has added a second measure on top of it and treated that addition as the line's true generative principle: a pulse, a beat, heard the way a listener follows a drum — an operation that depends on the listener, on the performance, on the ear brought to it, laid over a syllabic floor that lineage has always conceded but never treated as sufficient on its own. That addition has never been the poets' own. This book has tested that answer against eight independent corpora, four hundred and seventy lines of nineteenth- and twentieth-century poetry. It has failed every time. What has never failed, on the same evidence, is the answer Swahili-language poets themselves gave, in their own vocabulary, before any of this testing began: *vina na mizani* — the rhyme and the count.

How the test itself was built

None of what follows would carry weight if it rested on an impression rather than a procedure, and the procedure is worth stating before its results are. A claim earns the status of a scientific proposition, in Popper's sense that this book adopts throughout, only if it specifies in advance the exact observation that would defeat it. A theory patched after every contrary case, reclassifying each one as an exception or an error, cannot be defeated by anything and is not, for that reason, a theory of the kind that deserves belief. This is the standard applied here, to both sides of the argument at once and before either side was tested.

The application began with formalization, not with listening, and the formalization was itself a procedure the data were allowed to correct rather than a single act of stipulation. Chapter 8 translates the classical descriptions of *vina na mizani* into an explicit algebraic notation — a matrix for each meter, stating its syllable count and its rhyme variables as a formula, derived from reading how the poets themselves described their own practice rather than imposed on it. Each matrix was then checked against a corpus, refined wherever a corpus supplied a case the initial formulation had not anticipated, and checked again — an iterative interaction between the formal model and the data, in which the data held final authority throughout.

This is the reverse of what Werner did in the early 20th century (1917-20). She decreed, without a mapped corpus, without a stated method, and without argument, that a poetic line was measured by two beats rather than by syllables, and that decree has stood as an unrevised dogma against a century of poets' own testimony to the contrary. Where this book's formalism bent to what a corpus supplied, the accentual school's has bent every corpus to what it had already decided before examining one, a procedure this book names a prescriptive reductionism: the answer fixed in advance, and every case reduced to fit it regardless of what it actually contains. The same discipline was then turned on the theory this book contests. Vierke's *Grenzschema* and Minerba's trochaic innovation are not, in their own texts, stated with this precision — they are described, illustrated, asserted of a corpus in general terms. Chapter 2 does the work their own authors did not: it restates each claim as an explicit, falsifiable condition, the bipartite axiom as a count, the *Grenzschema* as a pair of fixed positions, before any of the eight conditions were run against a single recording. A theory is tested fairly only once it has been made as precise as its rival, and this book made it so before conceding it a single point.

Three independent readings were then taken, in a fixed order that none was permitted to revise once the next was seen: the first two on every line, the third on a sample. The first is deterministic: the printed text alone, and the ordinary rule of Swahili stress, yield a reading anyone can reproduce without a recording at all. The second is auditory: a trained ear, working from a formal recitation deposited alongside this study, screens that reading against what a listener actually hears. The third is instrumental, and it is the one this book insists on stating plainly rather than hiding behind a chart: fundamental frequency, intensity, and syllable duration, measured

and plotted from the recording itself, on what this study calls a hostile sample, composed for three distinct reasons rather than one. Some hemistiches were chosen specifically because they should be hardest for the accentual model to survive: the highest count of accentual groups a corpus offers, and the least symmetrical configuration among those that do satisfy its bipartite axiom. Others were drawn at random, an independent check against the suspicion that a sample selected entirely by the researcher's own hand might have been chosen to make the model fail. And a further few were chosen because they are the model's own ideal configuration realized without qualification, since a sample built only to attack a theory would not be a fair test of it. This instrumental layer is offered strictly as information, never as a way of extracting a meter from a signal — no meter is generated by an acoustic curve anywhere in this study — but it lets a skeptical reader verify, on an actual waveform, whether an annotated accent leaves the physical trace the ear reports. Every one of these curves, and the recording each is drawn from, and the script that drew it, is deposited for exactly this purpose: not to be taken on this book's word, but to be checked against the sound itself.

What was tested, and what followed from it

Two families of claim were placed under the same scrutiny, formalized in advance as eight conditions, so that neither could be judged by a looser standard than the other.

The first four conditions belong to the syllabic model — the claim, made by the Swahili tradition itself, that a classical line is generated entirely by its syllable count and its rhyme. Does the line's absolute length match what the form prescribes? It does, in the overwhelming majority of every corpus examined, and every departure is one the tradition itself names and licenses as *guni*, a defective verse — never an exception invented by this study to protect its own model. Does the rhyme lock into place in the strict syllabic form the tradition treats as primary, tested first in every case, with the looser vocalic match reserved for the rare, individually named line where the stricter form fails? It does, across corpora as different as a nineteenth-century prayer for rain and a twenty-first-century composition in which the Swahili language speaks in its own voice, the looser form invoked only where a specific line requires it and declared as such rather than assumed. Where a line is divided into segments, does that internal boundary hold — does a word ever straddle it? At every internal boundary tested in the *wimbo* and *shairi* corpora, never once. And

where a form chains its rhyme across an entire stanza, or across several stanzas in a recursive architecture, does that chain hold to its own rule? It does, term by term, including in *Mashairi ya Vidato*, the most demanding architecture examined in this book — a form in which the closing half-line of one stanza is repeated word for word as the opening half-line of the next, while its rhyme simultaneously migrates forward to seed the following stanza's own pattern. No accentual theory has an account of this structure, because no accentual theory has any means of noticing it: it is invisible to a model that listens for a beat, and entirely legible to one that reads the page.

The second four conditions belong to the accentual model — the claim, first asserted in the early 20th century (1917-20) and inherited without revision for a century, that beneath this syllabic surface a rhythmic pulse of two beats per line, or sometimes four, generates the true architecture of the line. Does a classical eight-syllable segment reliably divide into exactly two stressed groups? It does not: in every corpus tested, including the one composed by the tradition's own most celebrated master and offering the model its most favorable terrain, at least one segment in four contradicts the rule outright — and in four instances, in four poems of two meters and always in the same configuration, three separate and independent procedures, a deterministic reading of the text, a trained ear, and a physical instrument, agree that an eight-syllable segment carries three stressed groups, not two, a configuration the rule declares cannot exist. Do those two stresses, where they occur, fall on the specific syllables the theory calls its ideal pattern? At best, in the single most favorable corpus in this entire study, one segment in three; nowhere else does it clear one in four, and in one classical prayer for rain, built on segments of six and nine syllables rather than eight, the pattern has no syllabic frame to occupy at all — the theory is not merely wrong there, it has nothing to be wrong about. Does a genuine rhythmic foot recur, the way a drumbeat recurs? It does not, at either scale tested: the internal symmetry a foot requires is absent from the large majority of segments in every corpus, and the sequence of configurations from one line to the next is statistically indistinguishable from configurations drawn at random, never once showing the recurrence a rhythm would produce. And can the four-syllable foot this theory proposes as the shared engine of three different classical forms even be built from the syllable lengths those forms actually supply? Sheikh Ahmed Sheikh Nabhany's own *wimbo*, a single line cut by its rhyme chains into segments of six, three, four, and two syllables, admits no partition into tetrasyllabic

units by any arrangement whatever, and the near-totality of a fifteen-syllable *shairi* examined in this study excludes the foot the same way. The answer, in both, is that the foot cannot be constructed at all — not badly predicted, but arithmetically impossible, while continuing to be asserted as the generative law of the very corpus that excludes it.

One further test closes the account rather than reopening it. A separate control corpus of ordinary Swahili prose — composed for no metrical purpose whatsoever, read aloud under the identical protocol — realizes the accentual school's ideal configuration every time its ordinary words happen to supply two four-syllable units, eight times out of eight, and produces it once more inside a single word, from nothing but the stress a reader places on a tense marker. A reader alters the accentual shape of what she reads at the same rate whether she is reading a nineteenth-century epic or a sentence about the price of fish. What the accentual school has spent a century describing as the generative rhythm of Swahili verse is, on this evidence, a property of the Swahili lexicon itself — a statistical byproduct of how often ordinary words happen to be two or four syllables long — and not a property of the meter at all.

That this ideal is nonetheless achievable, given the right conditions, is demonstrated rather than merely argued: a text composed under a single constraint, that every word measure exactly four syllables, satisfies the accentual school's complete ideal on every one of its lines, without a single exception, while remaining, independently, an entirely well-formed classical stanza by the syllabic tradition's own rules. The demonstration cuts two ways at once. It proves the ideal is not an empty fantasy — it can be built, and this book has built it. And it shows that no classical poet, across four hundred and seventy lines of philologically established poetry, ever sustained it beyond a minority of lines, because sustaining it costs exactly what the tradition's actual lexicon — its compounds, its clitics, its words of two, three, five and six syllables — was never willing to surrender.

Where the error came from

This error is not a single misreading. It is a lineage, traceable from its first appearance to its most recent restatement, and every restatement inherits the same unexamined premise rather than testing it afresh.

Edward Steere, in 1870, calls the classical *utenzi* anapaestic — a term that, in its own classical home, names a pattern of syllable length, long against short. Swahili has no such length to measure, so what

Steere heard in the *utenzi*, whatever name he gave it, could only ever have been a matter of stress: the vocabulary is ancient Greek, but the judgment applied to Swahili, from this first instance onward, was already the modern one. Carl Büttner, in 1894, cannot settle the question he inherits: he classifies the same corpus as fluctuating between "iambic" and "trochaic," two more labels from the same Greek vocabulary, distinguished, in their modern sense, by which syllable in a pair carries the stress. That two of the earliest Western observers of the same body of poetry could not agree on which classical foot to impose on it is itself a finding: neither foot was there to find. Writing in 1917–20, Alice Werner asserts that the line is measured by two beats rather than by syllables — without a mapped corpus, without a stated method, without argument. A. M. Jones, working from a musicological angle in 1976, hears four beats where Werner heard two.

Clarissa Vierke, in 2011, builds an entire theoretical apparatus — she calls it the *Grenzschema*, adapting a term and a concept from Jakobson's work on an entirely different poetic tradition to an idealized, perfectly symmetrical pattern she proposes as the line's cognitive template — on Werner's figure of two (1917-20), passing an assertion made in the early 20th century into a doctoral thesis published nearly a century later with no intervening test of whether it was ever true. Two positions on the same eight-syllable line, both defended inside the very same school of thought, one insisting on two beats and the other on four, and no procedure exists anywhere in that school for choosing between them. A tradition that has examined one line for a century and a half and cannot agree on how many beats it contains has a problem no new terminology will resolve.

The prose controls of chapter 9 resolve the disagreement rather than taking a side in it. In ordinary sentences about food and visitors, the same reader produces major stresses on the third and seventh syllables and minor ones on the first and fifth: an analyst listening for two beats hears two, and one listening for four hears four, in the same utterance. Werner and Jones were both describing the Swahili language. Neither was describing its meter.

Emiliano Minerba, in 2025, inherits Vierke's *Grenzschema* without reopening it, and adds a further layer that only sharpens the pattern rather than breaking it. He first attempts to derive the Swahili line from medieval Yemeni poetic architecture, by way of a hypothesis that these forms traveled across the Indian Ocean; when that attempt runs into the fact that Swahili possesses none of the vowel-length

distinctions the Arabic and Persian metrical systems require, he does not abandon the trochee — the two-syllable, stress-then-slack unit at the heart of the whole tradition. He does not, for that reason, abandon the Arabic foot; he converts it. Where the Arabic system measures that foot by the duration of its syllables, long against short, Minerba reassigns its core a dynamic accent instead — a stress in place of a length the language does not have — and a foot measured this way, by loudness rather than duration, is by definition what a trochee is. He declares the result a "trochaic innovation," claiming it as the single shared engine behind three separate classical Swahili forms — a diffusionist hypothesis of considerable scope, asserted with as much confidence as brevity, in the space of five pages. The route changed once in a century and a half — Greek directly in 1894, Greek by way of Arabic in 2025 — but the destination and the operation did not: a duration the language does not mark is replaced, either time, by a stress the theorist supplies, and the resulting beat is called a trochee, in its modern sense, regardless of which language it was borrowed through on the way there.

One partial concession is owed to Minerba, and stating it plainly is more honest than omitting it: comparing a rhyme pattern across two cultures as a traveling motif is not, by itself, an unreasonable hypothesis — motifs travel. His error is the moment such a motif, compared at that modest level, is made to generate the measured line itself.

That leap does not survive contact with the data on its own, so it needs support the data cannot supply directly — support that comes from filtering what is allowed to count as a test of it. The trochaic, tetrasyllabic foot he needs comes free on a four-syllable segment, guaranteed by the word boundary and the language's own penultimate stress rule alone. On an eight-syllable segment it takes two such positions, the third and the seventh — a premise, that a segment carries exactly two stresses and no more, that Minerba never examines himself; it is Werner's unrevised figure (1917-20), passed through Vierke. Of the two positions, only the second comes free the same way, and even that proves less certain once measured against the recording, where the physical peak lands on the final syllable rather than the penultimate often enough to matter. The first position is not free at all — so the model widens instead, counting any odd-numbered position as the same confirmed rhythm and setting aside only the even ones. A target this loose, drawn to fit whatever the data supplies rather than tested against a fixed prediction, is the Texas Sharpshooter Fallacy in

its exact form — and it is drawn precisely this loosely because the diffusionist hypothesis it serves needs to hold across three meters at once, not one.

What this apparatus has tested, after all this refinement, is a narrow slice of classical Swahili verse.

The three layers, and the role the school confused

Classical Swahili verse separates cleanly into three layers, and the error traced above is, at bottom, a failure to keep them apart. Layer 1 is the syllabic law itself, *mizani* and *vina*, the count and the rhyme. Layer 2 is the architecture a composer builds on that law — the rupture matrix of a stanza, the recursive chaining of *Mashairi ya Vidato*. Layer 3 is the melodic overlay, *mahadhi*, added when the line is sung.

The tradition assigns these layers to two distinct roles, not one continuous process. The *mtunzi*, the poet-architect, works out the count and the rhyme entirely within Layers 1 and 2 — mentally, without pen or paper, or on the page, which fixes what the mind already resolved. Both are legitimate seats of the same structure; the opposition that matters is not the mind against the page, it is this invariant structure against what comes after it. What comes after it belongs to a second person entirely. The *mghani*, the singer, receives the finished composition and supplies Layer 3 alone, free to stretch a vowel, shift a tempo, or syncopate a phrase exactly because the *mtunzi*'s grid is fixed enough to survive that freedom without collapsing. The two roles form a contract, not a collaboration in real time: one builds what cannot break, the other performs what may vary as it likes.

Accent belongs to neither of these two builders. It is a property of the language itself, fixed by the ordinary rule that stress falls on a word's second-to-last syllable, external to the *mtunzi*'s architecture and to the *mghani*'s melody alike. Stated at its most direct: the accentual school's error is a category error, mistaking a property of the language for a property of the meter. Accent surfaces imperfectly whenever a line is read aloud, and it is bent further still once the *mghani* folds it into a tune. This is the precise shape of the century-old error, and it has its own name: a Performative Fallacy, a property of the language distorted in the one layer that was never meant to be measured for it, the *mghani*'s, mistaken for a law of the layer the *mtunzi* alone controls. Vierke's own corpus is, in all probability, drawn from exactly this confused layer: a chanted execution she does not herself claim follows a strict rhythmic pattern. A school that

never named these three layers, or the two roles that keep them apart, had no way of noticing it was doing this.

What that error looks like when named plainly

The pattern this century and a half of scholarship shares has a name, and this book does not shrink from using it. The stone ruins of Great Zimbabwe were credited to King Solomon's mines, to Phoenician traders, to a vanished white race — to anyone at all except the Shona people whose hands had laid the stone. Professional archaeologists showed in 1905, and again in 1929, that the Shona had built them. The colonial order refused to believe it for decades more. This book calls the pattern behind that refusal the Exo-African Postulate: the assumption, rarely stated and never argued for, that a sufficiently sophisticated achievement of African design must have its generative principle somewhere else. Swahili metrical scholarship has enacted the identical pattern on paper rather than in stone. When classical Swahili verse displays an architecture as demanding as *vina na mizani* — a single line divided by rhyme alone into four segments of six, three, four, and two syllables in succession — three of the four impossible to build from four-syllable units at all; a nineteenth-century prayer for rain, one of the most traditional compositions examined in this study, built throughout on lines of six and nine syllables rather than eight, offering the accentual model no segment of the length its entire apparatus requires, from its very first line; or a stanzaic chain, *Mashairi ya Vidato*, capable of the recursive complexity described above — the assumption has been, again and again, that such rigor could not simply belong to the language and the poets who speak it, and must instead be traced to Arabia, or to Greece.

The *Grenzschema* can now be named for exactly what it is. It is not merely an illustration of the Procrustean bed this book takes its title from. It is the bed itself: a fixed measure, invented elsewhere, onto which the living diversity of Swahili verse has been stretched where it fell short and cut down where it ran long, so that it would appear, on the page of a theory it never asked to be tested against, to fit.

What the language itself already refused to say

None of this required a discovery this book claims for itself. It required listening to a vocabulary that had already spoken, in terms of remarkable precision, and had simply not been heard.

Classical Swahili borrowed extensively from Arabic to describe its own poetry — *arudhi* for the science of metrics, *ubeti* for the stanza —

and it is precisely the fate of these borrowed words that gives the game away. In Arabic and in Persian alike, *ubeti* derives from *bayt*, which means literally a house, or a tent — the name given to a single line of verse, split into two mirrored halves built from feet measured by vowel length. In Swahili, that same house becomes a far larger one: an *ubeti* is almost never one line, but an entire stanza of four, five, or six lines, internally divided in ways that vary by form and owe nothing to a bipartite mirror. The Swahili-language poet kept the name of a house built for a single line and used it to house an entire stanza instead. The same holds for *arudhi* itself: adopted as a prestigious label, then evacuated of the quantitative substance the label originally carried, because Swahili draws no phonological distinction between long and short vowels at all, and a metrical system built on that distinction simply has nothing to attach to in a language that lacks it. Borrowing a label is not, by itself, the failure: *mizani* is borrowed too, from the Arabic word for a balance, a scale, and it names a mechanism that works exactly as its label suggests, a count that balances one line against the next. What fails is borrowing a label for an engine the language cannot run — and a scale, unlike a foot measured in vowel length, asks nothing of Swahili phonology that Swahili phonology cannot give it.

Kina, the rhyme, carries an ambiguity of its own in its very name. Sacleux traces it to two possible roots at once, and does not resolve between them: an Arabic term for the meter of a scanned line, or a Bantu word for a pit, a hole, the depth of water. One reflection follows from setting that unresolved root beside a fact that needs no resolving: *bahari*, the word this same tradition uses for a meter, is the same word Swahili uses for the ocean. Nothing in the record settles whether the poets themselves ever intended this double resonance between the ocean's depth and the meter's rhyme. But a surface can be borrowed from anywhere and still answer to what lies beneath it, and on the evidence this book has gathered, what lay beneath the classical Swahili line, however its surface was named, was never anything but its own depth.

The clearest proof of all is not a term the tradition redefined under pressure. It is a term the tradition never had reason to coin. Swahili absorbed the vocabulary of classical Arabic prosody with evident enthusiasm — *arudhi*, *ubeti*, *mizani* among it, the count, the rhyme, and the science of verse itself all carrying borrowed names — and it never absorbed or coined a single word for the foot, or for a rhythmic beat, though nothing in Swahili phonology would have resisted either:

a clean *tafila* would have assimilated as easily as any of the terms just named, and *taf'ila* (تَفْعِيلَة) itself simply never crossed over. The gap runs deeper than a missing positive term. A line that fails its syllable count or breaks its rhyme has a name in every treatise this book has examined, *guni*, a formal defect, stated as such without appeal; no line in the entire literature is ever called defective for missing a beat, because no concept of a missing beat exists to be violated. A tradition capable of naming its own errors with this much precision, and which never once reached for this particular error, was not failing to notice a rhythm. It was describing, faithfully, a system that never had one.

What was owed, and to whom

This book restores a description rather than proposing a new one, and the description was given, long before this study began, by the people best positioned to give it. Sheikh Kaluta Amri Abedi opened his 1954 treatise by insisting that a composer's thought must generate the rhyme, never the reverse, and gave the arithmetic of *vina na mizani* its first systematic statement. Shaaban Robert elevated that same architecture to the register of an ontological necessity: he opens his own treatise on the poetic art by comparing verse to the Nile, the intellectual lifeblood of the Swahili world, and defines the form itself as nothing but a measure in the pronunciation of syllables — eight by eight, sixteen by sixteen, seven by eight, in his own listed variants — with no mention anywhere in his taxonomy of a musical beat. His definition of a defective line names only a failure of that measure or of that rhyme, never of a rhythm. Ibrahim Noor Shariff gave the architecture its most exacting statement in *Tungo Zetu*, defining the syllable count as, in his own words, the count of the sounds in every word. Ahmed Sheikh Nabhany cataloged the classical system's full extent, thirteen recognized forms in all, of which this book examines three.

These are the voices of scholars and poets who wrote directly about the architecture of their art. A century earlier, one poet had already shown what that architecture could do rather than described it. Muyaka bin Haji, in the *Diwani* that carries his name, built in a single poem a chained rhyme structure — *Mashairi ya Vidato* — whose recursive logic this book has verified line by line and found without a single break across three separate stanza transitions.

Two Western scholars deserve to be named separately, because the usual account of this history erases the distinction they represent. Lyndon Harries described the classical Swahili line as answerable to

its own formal logic and to no other — sui generis, in his own term — and Jan Knappert anchored his own analyses in *vina na mizani* without ever feeling the need to superimpose a foreign stress pattern on top of it. Both men have been criticized, and not without grounds, for an Orientalist bias in their broader scholarship — for privileging the Islamic and Arabo-Persian strands of Swahili civilization over its Bantu foundations. This book takes that criticism seriously rather than setting it aside, and finds, on the specific question of meter, that it cuts against the very conclusion it might seem to support. Scholars inclined by their own stated bias to favor an Arabic frame had every reason to import the Arabic metrical foot wholesale, and did not. Confronted with the physical reality of the manuscripts, they described exactly what the manuscripts contained. That their fidelity to the syllabic evidence survived a bias that pointed the opposite direction makes that fidelity a stronger witness, not a weaker one.

What operated against these voices, and against the Swahili-language tradition more broadly, was rarely outright erasure. Both Vierke and Minerba cite Knappert and Harries by name, and cite them accurately. What operated instead was narrower and, for that reason, harder to see: within the same pages that correctly restate the syllabic reality, that reality is set aside for the search that follows, treated as settled and self-evident background rather than as the generative principle it is. Minerba himself concedes the syllabic count's authority over the classical line more than once, and then proceeds as though that concession changed nothing. A structure seen and immediately looked past is not a structure refuted. It is a structure this book insists be looked at again, and this time, held.

One further caution belongs here, because this book has no wish to be misread as the mirror image of what it criticizes. To reject an imported metrical framework is not, by itself, to adopt any particular political posture, and this study is not one. Nor does the critique advanced here indict Western scholarship as a whole, or Islamic and Arab influence on the Swahili coast as such — a civilization built on centuries of Indian Ocean exchange has nothing to apologize for in that exchange, and this book has drawn gratefully on French, German, and English philology throughout. The target is narrower and more precise: a specific theoretical lineage that mistook a borrowed vocabulary for a borrowed architecture, and, at the level this book has tested directly, mistook a property of the language, its accentuation, for the generative principle of the meter — against the meter itself, and against the Swahili-language poets who built it. That lineage tested

this mistake against remarkably little data for remarkably long, and can now be shown, on data it never previously faced, to have been wrong.

What remains for the reader who checks

Every recording examined in this study, every syllable-by-syllable measurement, every script that produced a figure from that measurement, is deposited in open access, alongside the data it processed. A theory shielded from this kind of scrutiny is precisely the condition that allowed an unexamined figure, asserted by Werner (1917–20), to pass into a doctoral thesis in 2011 without a single intervening test — and it is the condition this book refuses to extend to its own claims. What is owed, in place of a framework imported and never verified, is not another unverified framework offered in its place. It is a method that could not have produced this particular error twice, because every step of it can be checked by anyone willing to listen to the recordings and read the code. That is what Open Science means here: not a slogan appended to a finding, but the condition under which the finding was allowed to stand.

The classical Swahili line never needed a foreign engine to generate what this book has measured. *Vina na mizani* — the rhyme and the count — said everything the line required saying, before this study began. What this book has added is not a confirmation that description never needed. It is the falsification, on the imported model's own terms, of every claim that said otherwise.

Appendices

The textual materials reproduced in these appendices constitute the empirical corpus utilized for the three-tiered metrical and acoustic analyses in Chapter 9. Specifically, these source texts serve as the foundation for the methodological funnel applied in this study: (1) the deterministic computation of lexical stress (G-lex) based on the macro-Swahili penultimate rule, (2) the systematic auditory mapping of vocalized prominences (G-aud), and (3) the instrumental acoustic analysis utilizing Praat.

To ensure absolute methodological transparency, the corresponding audio recordings of these reading executions, alongside the digital source texts read by the L1 Swahili speakers, have been permanently deposited in open-access repositories via Zenodo. For comprehensive access details, including direct DOIs, please refer to Section 9.2.7.

A.1. Classical Swahili Corpus: Source Texts Provided to Fatuma Shaaban

1. *Utenzi wa Rasi 'Lghuli*

Awali bisumi 'llahi,
Jina la Mola 'llahi,
Pweke asiye shabihi,
Ndiye wahidi Kahari.

Arahamani Rahimi,
Ndiye hai 'lKayumi,
Mwenye ezi ya dawami,
Alipendalo hujiri.

Huwa Lahu 'lJalali,
Ndiye Muizi Mudhili,
Ndiye Hakimu Adili,
Ndiye Mjuvi wa siri.

Huwa 'lLahu Manani,
Mtukufu mwenye shani,
Khaliku alinsani,
Bi unasi wa dhukuri.

Huwa 'lLahu Wabari,
'Wahidi 'lKabari,
'lKabiri 'lGhafari,
Mudabiri li umuri.

Ndiye Karibu mjibu,
Ndiye Ghafiri dhunubu,
Ndiye Satari aibu,
Ndiye Rahimu shukuri.

Huwa Lahu 'lkhaliki,
Ndiye Muumba riziki,
Kuweneza makhuluki,
Wakuu hata saghiri.

Sharaka wa gharibia,
Suheli wa kibilia,
Kaweneza wote pia,
Wale bara na bahari.

Alla Alla walimwengu,
Tumhimidi Muungu,
Muumba inti na mbingu,
Majabali na shajari.

Tumhimidi Jalali,
Ni jawabu afadhali,
Kumsalia Rasuli,
Ndiye jawabu khiari.

Tumsalie Msharafu,
Tume wa Mola Latifu,
Kesho tupate wokofu,
Twepukane na sairi.

Tumsalie Habibu,
Asa ikawa nasibu,
Kwepukana na adhabu,
Hiyo siku ya nushuri.

Nikomile kuradidi,
Nimshukuru Wadudi,
Na sifa za Muhamadi,
Kumsifu sikadiri.

Ewe mwanangu Khamisi,
Nipatie karatasi,
Uniletee upesi,
Pasiwe kutaakhari.

Karatasi iwe shamu,
Iliyo njema naimu,
Na pambo ni maadimu,
Twaa hiyo ya hariri.

Wendapo ukinunua,
Jitahidi kuteua,
Nawe ni mtu mwelewa,
Si mchache wa nadhari.

Unapokwenda dukani,
Usiusiri ndiani,
Ufike kwa Selemani,
Au kwa Isa mzuri.

Mpatie na mapesa,
Uyatukue kabisa,
Ukiwasili kwa Isa,
Usiwe kutaasari.

Chamba huyo Selemani,
Hukumuona dukani,
Na urudi iyo hini,
Ende kwa Isa mzuri.

Dukani utaporudi,
Na safidiye madadi,
Na nduguyo Masuudi,
Ogopa kumudabiri.

Utaporudi dukani,
Ukiwasili nyumbani,
Usiwe kutamakani,
Usafidi zingifuri.

Kwa papo uyateue,
Zayobaki uyatoe,
Usafi hata ujue,
Itiwapo inawiri.

Unipatie kibao,
Kizuri kisio tao,
Na uzi na uje nao,
Ulio mwema khiyari.

Na uzi uwe na tasi,
Upote ungie kasi,
Hipigwa iwe nafasi,
Ituliye mistari.

Pana hadithi ajabu,
Yandikwa ndani kitabu,
Kwa lugha ya waarabu,
Nayo hadithi khiyari.

Hadithi hiyo yuani,
Ina maneno bayani,
Ya vita vya 'lYamani,
Zamani zake Bashiri.

Haiona haisoma,
Maana ni kufahama,
Hatibu wangu mtima,
Huwaza na kufikiri.

Hufikiri nikiweza,
Haona kunipendeza,
Moyoni nikilekeza,
Kutunga kwa ushairi.

Hapenda kuibadili,
Kwa lugha ya suwahili,
Kiarabu ni kalili,
Wajuao tafisiri.

Nikawa kujikalifu,
Kiarabu hawakifu,
Maneno hayasanifu,
Kuwapa zote khabari.

Nachelea ni mjinga,
Sijui sana kutunga,
Na maneno kuyapanga,
Yataka mtu mahiri.

Wala mi si mualimu,
Wala sijui ilimu,
Nataka kutaya mamu,
Zidini kuniudhuri.

Msije kuona kina,
Kisichoshika maana,
Mkiwamo kuninena,
Wakuu hata saghiri.

Mkanandika aibu,
Wala pasipo sababu,
Kunirudi ni thawabu,
Wala nami sikasiri.

Ewe mwanangu Khamisi,
Shika hima karatasi,
Na kalamu kwa upesi,
Pasiwe na kuusiri.

Ogope kutakarabu,
Nikwambie matulubu,
Nikupalo ukutubu,
Usiwe kutaakhari.

Hima andika upesi,
Suturi ziwe nafasi,
Na maneno ukakisi,
Usandike mshadhari.

Na khati iwe fasiha,
Maneno yawe wadhiha,
Bi udhuma na fataha,
Na jazima na kasiri.

Iwapo uzifanidi,
Kwa mada na tashihidi,
Ziwe baidi baidi,
Zisiwe mudaathari.

Na nukuta tia pia,
Usiwate hata moya,
Uzipange kwa sawia,
Zisiwe kutaghayari.

(Mgeni bin Faqihi, 1979, pp. 1-2)

2. *Utendi wa Haudaji*

Bisimillahi awali,
Isimu yakwe Jalali,
Ndrio mwando wa kauli,
Wa mt^{h}u kutamkia.

Arrahmani, Azizi,
Rabi ndriye Muawazi,
Wema, wawi huwajazi,
Ulimwenguni Jalia.

Na Arahimi karima,
Ni mteuzi wa wema,
Kesho janati nakima,
Mola taiteulia.

Alihamdu lillahi,
Kula mwenye kumadihi,
Ndakwe, asiyoshabihi,
Mola wa ziumbe pia.

Sifa zot^{h}e njema hizi,
Zeneo kama mwezi,
Ni zakwe yeye Azizi,
Hazikadiri mmoya.

Na kutosha ziadaye,
Ni rehema yakwe yeye,
Mwenye ezi Mungu ndriye,
Ambao kaonekea.

Na wokofuwe kamili,
Humwendrea murisali,
Na sahabaze na Ali,
Milele huwashukia.

Ulinenele hashima,
Mtume wakwe Karima,
Wokofuwe na rehema,
Wa Mungu hutuilia.

Tʰumwa wetʰu mfadhali,
Ndriye muumbwa awali,
Kapowa na manazili,
Yasiyomtindikia.

Hifadhi tuzie kwenu,
Ni mengi, ni nyingi mno,
Ndripo kʰakosa maneno,
Siziwezi kuzitaya.

Dibaji nimekhitimu,
Kwa cheo na takadhimu,
Sasa tʰaka kunudhumu,
Nimezoyakusudia.

Hadithi nimeiona,
Ya Muhamadi amina,
Yalipoingia fitina,
Yakwe na ujuhalia.

Na fitina zikikwaka,
Zikatupa na mipaka,
Alihajiri Maka,
Mtume wetʰu nabia.

Maka alipohajiri,
Kwendra Madina bashiri,
Wenyewe walishakiri,
Kwa tʰumwa kuwangilia.

Alitoka ye mpweke,
Hakutwaa taa mbake,
Fatuma na wake wake,
Maka wakaikalia.

Wake wotʰe na zijana,
Hakwendra nao amina,
Na utendi lakwe ina,
Ya haudaji sikia.

Utendi mbwa haudaji,
Ndrilo inale, wambuji,
Zalizongia siraji,
Muhammadi kumwendrea.

Tʰumwa kule Madinati,
Alipokwisa kuketi,
Aliangata sauti,
Kamwamkuwa Alia.

Kamba: "Bunu amu yangu,
Kurubia hapa pangu,
Ninene maneno yangu,
Upate kuyasikia."

Akajibu Haidari,
"Upendalo, nikhubiri,
Ambalo walikhitari,
Siche kunikarifia."

Akatamka amini,
Kamba Aba lHasani,
"Mno nimewatamani,
Wake wangu na dhuria.

Usiku silali tena,
Sioni raha mtana,
Mi nao kutoonana,
Mno yamenikulia.

Nililo nalo, kʰwambie,
Enendra kanitwalie,
Moyo wangu upumbae,
Mato kʰiwaangalia."

Akajibu Haidari,
Kamba: "Ewe mukhutari,
Kutumika nimekiri,
Thakwendra kukutwalia.

Sasa hiyau tʰatoka,
Tʰumwa wangu, nendre Maka,
Hao umezowataka,
Inshallah, tʰakwetea."

Hayo alipopulika,
Alii mezotamuka,
Moyo walimkunduka,
Na raha ikamngia.

Kafurahi muungamu,
Kamba: "Ewe, bunu amu,
Thakupa na mare makamu,
Waraka wa afia.

Waraka mwenye ziasi,
Wendre nao, mmarisi,
Kwa baba wetʰu upesi,
Bunu Abi Talibia.

Thakwandikia waraka,
Wendre nao hata Maka,
Kwa baba wetʰu, pulika,
Abasi, mwenye mafaa.

Na waraka hono ndrani,
Nitiziezo yakini,
Ni wakia khamusini,
Moya haikupunguwa.

Wakia hizo, sahibu,
Khamusini kwa hesabu,
Nda fedhati na dhahabu,
Amuyo kumnikʰia.

Basi uye, ujilisi,
Uishike karatasi,
Nikweleze kwa ziasi,
Kula yambo moyamoya.

Waraka hunu nandika,
Ili kuupeka Maka,
Wapate kuyapulika,
Maneno, tʰakokwambia.

Bismillahi radidi,
Isimu yakwe Wadudi,
Ndrimi tʰumwa Muhammadi,
Mwenye kuwashufakia.

Waraka utoka kwangu,
Mimi, mtume wa Mungu,
Kuwakhofisha wazungu,
Na mema kuwakedhea.

Utoka kwangu Hashima,
Kuwabashiria mema,
Mumfuwate Karima,
Kesho mpate kukia.

Hiyo ndriyo yangu ezi,
Nipozewa ni Azizi,
Kuwa mimi muombezi,
Umati kuwaombea.

Mfuwateni Manani,
Nami, mtume amini,
Kesho mwingie pʰeponi,
Huri mkae pamoya.

Fuwatani, ndrilo pato,
Mfahamu, kuna nyoto,
Huwaka kwa mivukuto,
Hauna kuzizimia.

Nyut^{h}e nimewakhubiri,
Na kuwakhofisha nari,
Muwate kutakabari,
Mkiri Mola Jalia.

(Vierke, 2011, pp. 479-489)

3. Nyimbo

3.a. Wimbo wa miti

Yana, niwene kisa adhimu,
Tena, tafusirini walimu,
Miti yalikipijana, mtendeti na mdimu.

Hima, ukaya mbiyo mpwera,
Nyuma, muembe una hasira,
Mzabibu ukisema, kunyamaa ni ujura.

Mara, ukadhihiri mtesi,
Sura, zina ushiu na kasi,
Ukiwambiya mpwera, kwetu hakuna rakhisi.

Papo, ukaya mpilipili,
Hapo, ukinena kwa ukali,
Ukiapa kwa ziyapo, miti pia siijali.

Punde, wasiyesa kani zao,
Nde, wawene mti uyao,
Nao ni mberemende, una panga na mangao.

Zita, hatuchi twapo uwawa,
Uta, na zembe tumetukuwa,
Mbuyu ulipopita, ukasimama kwa muwa.

Hela, watani yenu mayowe,
Tela, bure asiwazuzuwe,
Kuna mti una hila, humshinda kitunguwe.

Shina, la mambo haya nayuwa,
Sana, nimeziye kutambuwa,
Muyungu ndio fitina, waloalisha mauwa.

(Mbele, 1996, p. 72)

3.b. Sina Hali

Sina hali, moyo wangu nidengene,
Sihimili, wala sipendi mngine,
Tafadhali, nakuomba tuonane.

(Shariff, 1988, p. 46)

3.c. Kukwepuka nana

Kukwepuka nana, sitaki, nasonona, ndani.
Kusubiri tena, ni dhiki, muhibana, lini.
Nyonda kuonana, ashiki, hali sina, shani.

(Shariff, 1988, p. 48)

4. *Waji Waji. Dua ya Kuombea Mvua*

Ai subuhana, Mola wetu shirika,
Ndiwe Rahamana, muwawazi usio shaka,
Warehemu wana, na wazee wenye mashaka,
Waijuwa sana, hali yetu kati chaka,
Juwa la mtana, kula siku lisilokoma.

Tu waja dhaifu, mlangoni pako tuemele,
Kwa nia takafu, shida yetu Rabi yondole,
Hatuna wokofu, mtatuzi uyatatule,
Kwako ni khafifu, ni mafupi yangawa male,
Mwingi wa lutufu, muwawazi wawi na wema.

Rabi tuawaze, tungaasi kwa rehemayo,
Situpatilize, kwa makosa tuyakosayo,
Tupe barakaze, Muhamadi ndiye tumwayo,
Na wote nduuze, waombezi wa mambo sayo,
Utunyesheleze, vuwa nyingi yenye rehema.

Tuletee vuwa, tuondolee nyota na juwa,
Ndiwe Mwenye kuwa, ufishao na kufufuwa,
Hali waijuwa, huhitaji kuarifiwa,
Hasha kuemewa, ni amuri Kuni ikawa,
Wala kutendewa, Jala Rabi mwenye adhama.

Wangi waayisi, kwa mateso kuwa tawili,
Wana wasiwasi, kwa madhambi yao thakili,
Rabi tunafisi, muungwana uso mithili,
Mwingi wa ukwasi, tupe naswi utujamili,
Na yawe mepesi, kwa Kuni yako kalima.

Saya tutakayo, hako mtu katu muweza,
Wala mwenye nayo, ela Wewe mwenye maiza,
Na tutamaniyo, kutendewa ni yetu aza,
Kwa yote twambayo, tutendee ili majaza,
Tushukuru kwayo, ifurahi yetu mitima.

Muomba Wadudi, havi shaka twaa yakini,
Apate muradi, na matako yakwe moyoni,
Pindi akinadi, ja Ayubu Rabi masani,
Katu hamrudi, mja wake mwenye maani,
Hata Namrudi, kafirimbi asosilima.

Kwa yote masimu, twakuomba Mola Karimu,
Kwanza ni Adamu, Mtumeo na muungamu,
Na wote Kiramu, wandikao kula kalamu,
Loho na kalamu, na zuozo zote kadimu,
Na wenye masimu, mawalii na waja wema.

Rabi tutubie, ya nusuha yenye nadama,
Utughufurie, kuu dhambi iwe lawama,
Nawe uenee, Uduuni yako kalima,
Tutawasilie, kwa Isimu yako adhama,
Dua naipae, na kabuli ije kwa hima.

Dua na itimu, kwa hamdi na taadhima,
Na iwe naamu, na kabuli ije kwa hima,
Wapate naimu, na ajiri wenye kusoma,
Sala na salamu, na mbawazi zendee tumwa,
Na wote kiramu, sahabaze wenye dhimama.

Rabi takabali, na kabuli iwe ajili,
Yote tuyanali, tutakayo kwa tasihili,
Hatuna amali, tutendayo yenye kabuli,
Ela kwa ajili, na rehema zako Jalali,
Ndiyo tawasuli, maombezi yetu wafama.

Twataka amani, ya ghadhabu zako Jabari,
Huku duniyani, na Kiyama siku ya nari,
Tutie peponi, tusakini katika dari,
Utupe hisani, wake wema na kukhitari,
Nao wildani, watupambe mbele na nyuma.

Na mwenye kuomba, kutendewa ni yakwe jaza,
Hata mwenye kwimba, jaza yakwe yawa kutuzwa,
Wallahi 'mi namba, aombaye Mola Aziza,
Kula matilaba, yuwapewa na nyingi nema,
Na tutamaniyo, kutendewa ni yetu aza.

Nami ningatunga, siyajui ni kutamani,
Sijui kupanga, lulu ngema na marijani,
Hatauwa shanga, makuruto yaso thamani,
Kwa nyote malenga, muonapo penye lahani,
Hamuwi wajinga, mufanyapo mwapata mema.

(Chiraghdin & Nabhany, 1987, pp. 45-47)

5. *Uwongo*

Karatasi nakutuma, leo unitumikiye,
Kabalegheshe kalima, uzidi kusema naye,
Kudanganyana si kwema, kulla jambo kwa ndiaye,
Yalekeze yalekeye, tus'ambiane uwongo.

Yalekeze yalekeye, waolezi wayaone,
Tujivune khatimaye, mitima yetu isone,
Hakuna liso nduuye, mambo pia yafenene,
Sitaki tudanganyane, na kwambiana uwongo.

Sitaki tudanganyane, ndipo haya nakwambia,
La kweli tuambiane, zitakate zetu nia,
Na ufanyapo mangine, mwisho wakwe ni vibaya,
Vitakuja kutukia, tus'ambiane uwongo.

Vitakuja kutukia, uyaone matukuzi,
Na kwamba wapendelea, n'elezato waziwazi,
Nipate yafahamia, nifunue masikizi,
Tuuwate upuuzi, tus'ambiane uwongo.

(Hichens, 1940, pp. 92-93)

A.2. Contemporary Corpus: Kiswahili!! (Voiced by the Author, Ustadh Mau)

The written text from which the poet read in Lamu in 2005, as first published in Roy (2013b, pp. 48–53). One misprint is corrected (3.3, *ziyongonzi* for *ziyongozi*), and aspiration, written th in seven words, is set as a superscript h. Where the poet's reading departs from this text, the analysis follows the reading; the divergences are listed in section 9.6.2.

Kunyamaa nimechoka, t^{h}anyamaa hata lini?
Wanangu huniepuka, kuwaona natamani
Walobaki kunishika, si wangu ni wawendani
Mimi nimewatendani, mbona mwanipija zit^{h}a?

Wanangu mimi wa damu, wana wa Uswahilini
Asili hawana hamu, Yakuniyuwa ni nani?
Wamenatiya kaumu, na wana wa majirani
Kosa langu kosa gani? mbona hunipija zit^{h}a?

Mimi mamenu si t^{h}asa, wala sina punguwani
Nimezaa Wamambasa, na kungine Zisiwani
Nizee wanasiyasa, na ziyongozi wa dini
Mafundi wa kulla fani, na washujaa wa zita

Ndimi mamake Muyaka, piya Mwengo Athumani
Nazahidi kadhalika, na wengi wake wendani
Ali Koti na Mataka, wote mbwa moya karini
Walitoka matumboni, wakawaa kama nyota

Inkishafi ngaliya, ukisome na Kidani
Ndipo takapokuelea, nikwambiyao mwendani
Ni t^{h}ungo zimebakiya, na hazifia athilani
Walozit^{h}unga ni nyani? Ni wanangu walopita

Na malenga wa Mvita, na piya Chiragudini
Nyayo walizifuwata, hawakukiri uduni
Nabahany huteta, lakini hufaliyani?
Ndiye pweke uwandani, ingawa ameikita

Bado kuzaa naweza, siyakoma ukingoni
Lakini mumenipuza, mumeitowa fuwoni
Wangine meitokeza, kunipangiya kanuni
Musamiyati kubuni, nyinyi mulipowata

Huliya kisikitika, changaliya jaridani
Wengi wanaoandika, si wanangu ni wageni
Idhaani kadhalika, wapeka t^{h}ungo ni nyani?
Wengi hawatoki pwani, licha kuwa mbwa Mvita

Angaliya na zitabu, zisomeshwao shuleni
Hazandikwi na Rajabu, Si Sudi wala si Shani
Njoroge ndiyo katibu, ashishiyeo sukani
Charo na wake wendani, nao nyuma hufuata

Hualikwa kongamano, chendra hurudi ndiyani
Huona utungu mno, kuwa nyinyi siwaoni
Na huziuma zitano, lakini nitende nini?
Wanangu mumeikhini, mamenu mumeniwata

Nahuliya kwa matozi, changaliya mitihani
Wanafundi wa Kibwezi na wa Kisumu ziwani
Ndiyo wanaobarizi, waliyoko kileleni
Mulotoka kwetu pwani, muko tini hukokota

Wafanyao utafiti, wa digrii zuwoni
Waswahili ni katiti, au hawapatikani
Ni nyani nimlaiti? Mwenye makosa ni nyani?
Mimi hamunithamini, mgine hamukupata

Kiwasikiya hunena, huniongonga moyoni
Tharafu hakuna tena, nahau naitamani
Na hata ladha hayana, kama mashapu kanwani
Sielewi hunenani, huimba au huteta

Lau Muyaka tarudi, ae tena duniyani
Mwanangu itambidi, kwenenda mahakamani
Ahe tena mashahidi, waniyuwao yakini
Nyute mwende gerezani, kwa hatiya kuwapata

Wallahi hamuna ghera, wala hamuna imani
Hamuna la kuwakera, kuwa hamunithamini
Mimi ni kama mpwira, hutezewa uwandani
Hupijwa teke ndiyani, nakula anaepita

Hata kwenye ushairi, waso wangu wamebuni
Zilizo huru bahari, kwa kuoleza wageni
Mimi hayo siyakiri, si mashairi kifani
Hayo yote ni kwa nini? Hizo ni mbinu za zita

Hambiwa mwenyewe sina, hini ni ajabu gani?
Huwae kakosa shina? Kawa na tandu yangani?
Nyani alonipa ina? Alonandika ni nyani?
Kiwa si uswahilini, ni wapi nilipopata?

Kuwa wengi huninena, si dhalili athilani
Ya kuwa wenyewe sina, Kiingereza hamuoni?
Hunenwa na wengi sana, pembe zote duniyani
Kina na kwao shinani, miziye haikukata

A.3. Prose Control Reading Text: Source Material Provided to Fatuma Shaaban

1. Barua yake ni ndefu.
2. Wamesahau kalamu.
3. Mwalimu wangu ni Juma.
4. Juzi tulikutanika.
5. Yeye ameleta maji
6. Tulikuwa tumechoka.
7. Safari ya siku tatu.
8. Tulikutanika jana.
9. Kalamu hizi ni zangu.
10. Dirisha limevunjika.
11. Barua, kalamu, meza.
12. Tumeandika barua.
13. Nyumba yetu imejengwa.
14. Sasa wanaelewana.
15. Nyumba yetu ina moto.
16. Hatukuwa tumelala.

17. Jana walipika chai.
18. Walipokea barua.
19. Watu wote wamechoka.
20. Amenunua mikate.
21. Tulipanda basi jana.
22. Gari lake lina rangi.
23. Wote wanajitahidi.
24. Mama yake amefika.
25. Mwalimu anafundisha.
26. Habari za leo hizi.
27. Kalamu imepotea.
28. Anafanya kazi ngumu.
29. Nilikuwa nimechoka.
30. Mkate huu ni wako.
31. Mimi ninawaelewa.
32. Samaki, mkate, maji.
33. Dawa hii ina ladha
34. Mwalimu wa shule yetu.
35. Ndugu yangu ana homa.
36. Kitabu hiki ni chake.
37. Wanafunzi wanasoma.
38. Baba yangu ana gari.
39. Leo barua ni yake.
40. Walikuwa wanaimba.
41. Kesho safari ni ndefu.
42. Anataka kusafiri.
43. Dada yangu anasoma.
44. Dada yake ana kazi.
45. Wageni wetu nyumbani.
46. Ninawaelewa wote.
47. Juzi habari ni mbaya.
48. Jana walikutanika.
49. Watu wote wana njaa.
50. Wote wamekaa chini.
51. Mama huyu ana pesa.
52. Walikuwa hapa jana.
53. Amefungua dirisha.
54. Kaka yake anacheza.
55. Mgeni ameondoka.
56. Sasa samaki ni ghali.
57. Magari yamesimama.

58. Tumeweka pesa benki.
59. Mvulana amelala.
60. Hatukuwapelekea.
61. Nimeona gari lake.
62. Hawakusikia vyema.
63. Hawakuwaelewesha.
64. Habari zimeenea.
65. Kesho tutasoma sana.
66. Alinunua mkate.
67. Baba yangu amelala.
68. Mnataka kuondoka.
69. Hapa mvua ni nyingi.
70. Wamejitahidi sana.
71. Hapa chakula ni chetu.
72. Leo wamejitahidi.
73. Hapa tunapika chai.
74. Tunatafuta samaki.
75. Na hawakutuelewa.
76. Leo tumeanza kazi.
77. Wageni wamewasili.
78. Wamekaa hapa sana.
79. Kaka yangu ana duka.
80. Dirisha la nyumba yetu.

A.4. Marked-Syntax Control Corpus

1. Nyama yote hula paka.
2. Wema sana watu hawa.
3. Tunda hili yeye hula.
4. Maji haya kunywa sasa.
5. Kazi zote fanya wewe.
6. Pesa zetu watu hula.
7. Chakula hiki kaleta.
8. Kitabu chako nasoma.
9. Matamu sana matunda.
10. Kisu hiki nimetwaa.
11. Gari letu wameiba.
12. Hapa wamefika watu.
13. Kazi wanafanya sasa.

14. Duka letu amefunga.
15. Sana wanapenda nyama.
16. Wazuri wageni hawa.
17. Nyumbani kwetu kalala.
18. Hatutawakaribisha.
19. Alipotufahamisha.
20. Waliyotufundishia.

References

Abdulaziz, M. H. (1979). *Muyaka: 19th century Swahili popular poetry*. Kenya Literature Bureau.

Abdulkadir, M. A., & Frankl, P. J. L. (2013). 'Kiswahili': A poem by Mahmoud Ahmad Abdulkadir to which is appended a list of the poet's compositions in verse. *Swahili Forum*, 20, 1-18. https://d-nb.info/1238600441/34

Abedi, K. A. (1954). *Sheria za kutunga mashairi na diwani ya Amri*. Eagle Press.

Allen, J. de V. (1981). Swahili book production. *Kenya Past and Present*, 13, 17-22.

Allen, J. W. T. (Ed.) (1971*). Tendi: Six examples of a Swahili Classical Verse Form with Translations & Notes*. Africana Pub. Corp.

Anyanwu, R.-J. (2001). *On the manifestation of stress in African languages*. Typology of African Prosodic Systems Workshop, Bielefeld University, Germany.

Ashton, E. O. (1962). *Swahili Grammar (Including Intonation)*. Longmans, Green and Co LTD. First edition 1944.

Balbiani, F. (2023). Mission – Kolonialismus – Nationalsozialismus. Ernst Dammann und die Hamburger Afrikanistik, 1930–1937 (Hamburger postkoloniale Studien 8). Allitera.

Baraza la Kiswahili la Taifa (BAKITA). (2022). *Kamusi kuu ya Kiswahili* (Toleo la 3). Longhorn Publishers PLC.

Berwouts, K. (1991). *Le sein de la mère. Introduction à la littérature classique et moderne en swahili*. Les cahiers du CEDAF. ASDOC Studies.

Bohas, G., & Paoli, B. (1993). Métrique arabe : une alternative au modèle xalilien. *Langue française*, (99), 97–106.

Brielle, E. S., Fleisher, J., Wynne-Jones, S., et al. (2023). Entwined African and Asian genetic roots of medieval peoples of the Swahili coast. *Nature*, 615, 866–873. https://doi.org/10.1038/s41586-023-05754-w

Brown, H. (1985). *History of Siyu: The development and decline of a Swahili town on the northern Kenya coast* (Unpublished doctoral dissertation). Indiana University.

Brown, H. (1988). Siyu: town of the craftsmen, a Swahili cultural centre in the eighteenth and nineteenth centuries. *Azania*, 23, 101-113.

Büttner, C. G. (1894). *Anthologie aus der Suaheli-Litteratur (Gedichte und Geschichten der Suaheli)*. Verlag von Emil Felber.
Chiraghdin, S., & Nabhany, A. S. (1987). *Malenga wa karne moja*. Longman Kenya.
Cornulier, B. de. (1982). *Théorie du vers : Rimbaud, Verlaine, Mallarmé*. Seuil.
Cornulier, B. de. (1995). *Art poétique : notions et problèmes de métrique*. Presses Universitaires de Lyon.
Diouf, J.-L. 2009. *Grammaire du wolof contemporain*. L'Harmattan.
Duhem, P. (1969). *To save the phenomena: An essay on the idea of physical theory from Plato to Galileo* (E. Doland & C. Maschler, Trans.). University of Chicago Press. (Original work published 1908)
Ehret, C. (1995). *Reconstructing Proto-Afroasiatic (Proto-Afrasian): Vowels, tone, consonants, and vocabulary*. University of California Press.
Encyclopédie de l'Islam (Vol. 2). (1961). E. J. Brill; G. P. Maisonneuve & Larose.
Fontein, J. (2006). *The Silence of Great Zimbabwe: Contested Landscapes and the Power of Heritage*. UCL Press.
Garlake, P. S. (1973). *Great Zimbabwe*. Thames and Hudson.
Goody, J. (1977). *The Domestication of the Savage Mind*. Cambridge University Press.
Greenberg, J. H. (1947). Swahili Prosody. *Journal of the American Oriental Society*, 67(1), 24-30.
Harries, L. (1962). *Swahili Poetry*. Clarendon Press.
Hartman, C. O. (1980). *Free verse: An essay on prosody*. Princeton University Press.
Hashil, H. S. (2009). *Kurunzi (The torch): Swahili poems by Hashil Seif Hashil*. Swahili Consult.
Hichens, W. (1939). *Al-Inkishafi. The soul's awakening*. Sheldon Press.
Hichens, W. (1940). *Diwani ya Muyaka bin Haji al-Ghassaniy*. Witwatersrand University Press.
Hyman, L. M. (2013). Penultimate lengthening in Bantu: Analysis and spread. UC Berkeley Phonology Lab Annual Report.
Jakobson, R. (1933). Über den Versbau der Serbokroatischen Volksepen: 1933. In *Proceedings of the International Congress of Phonetic Sciences*. First Meeting of the

Internationale Arbeitsgemeinschaft für Phonologie Amsterdam 3-8 July 1932 (pp. 44–53).
Johnson, F. (1939). *A standard Swahili-English dictionary* (Founded on Madan's Swahili-English dictionary). Oxford University Press.
Jones, A. M. (1976). Swahili epic poetry: A musical study. *African Music*, 5(4), 105–129.
Jungraithmayr, H. (2007). Ernst Dammann (1904–2003). *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 157(1), 1–6.
King'ei, K., & Kemoli, J. A. (2001). *Taaluma ya ushairi*. Foto Form Ltd.
Klein-Arendt, R. (1987). Kingozi: Extinct Swahili dialect or poetic jargon? A historical, dialectological and contextual analysis. *Sprache und Geschichte in Afrika*, 8.
Knappert, J. (1967). *Traditional Swahili Poetry: An Investigation into the Concepts of East African Islam as Reflected in the Utenzi Literature*. E. J. Brill.
Knappert, J. (1971). *Swahili Islamic Poetry* (Vol. 1-3). E. J. Brill.
Knappert, J. (1979). *Four Centuries of Swahili Verse: A Literary History and Anthology*. Heinemann Educational Books.
Knappert, J. (1983). *Epic Poetry in Swahili and other African Languages*. E. J. Brill.
Knappert, J., & Van Kessel, L. (2010). *Dictionary of literary Swahili*. The Edwin Mellen Press.
Knappert, J. (1999). *Grammar of literary Swahili*. The Edwin Mellen Press.
Kyamba, A. N. (2022). Mchango wa Mathias Mnyampala katika Maendeleo ya Ushairi wa Kiswahili. *Kioo cha Lugha*, 20(1), 130-149. https://doi.org/10.4314/kcl.v20i1.9
Leslau, W. (2004). *Concise Amharic dictionary: Amharic-English, English-Amharic*. University of California Press; Otto Harrassowitz.
Lodhi, A. Y. (2004). Strategies of emphasis and intensity in Swahili. *Africa & Asia*, *4*, 1–12.
Lord, A. B. (1960). *The singer of tales*. Harvard University Press.
Maw, J., & Kelly, J. (1975). *Intonation in Swahili*. School of Oriental and African Studies, University of London.
Mazrui, A. M., & Shariff, I. N. (1994). *The Swahili: Idiom and Identity of an African People*. Africa World Press.

Mbaabu, I. (1985). *New horizons in Kiswahili: A synthesis in development, research and literature*. Kenya Literature Bureau.

Mbele, J. L. (1996). Wimbo wa Miti: An Example of Swahili Women's Poetry. *African Languages and Cultures*, 9(1), 71–82. https://doi.org/10.1080/09544169608717800

Mgeni bin Faqihi. (1979). *Utenzi wa Rasi 'lGhuli* (L. Van Kessel, Ed.). Tanzania Publishing House.

Miehe, G. (1979). *Die Sprache der älteren Swahili-Dichtung (Phonologie und Morphologie)*. Dietrich Reimer.

Minerba, E. (2025). *Meters in Motion: History of Swahili and Wolof Prosodies*. De Gruyter.

Mnyampala, M. E. (1963). *Diwani ya Mnyampala*. East African Literature Bureau.

Mnyampala, M. E. (ca. 1965). *Utenzi wa Zaburi*. Ndanda Mission Press.

Mnyampala, M. E., Roy, M., & Mnyampala, C. M. (1966). Primary Archival Extracts: Mathias E. Mnyampala's Unpublished "Mashairi ya Vidato" Manuscript (1966-1967). Zenodo. https://doi.org/10.5281/zenodo.21362077

Mnyampala, M. E. (2011). *Historia ya Hayati Sheikh Kaluta Amri Abedi (1924-1964)*. Ahmadiyya Printing Press.

Mnyampala, M. E. (2013a). *Maisha ni kugharimia* (Kimehaririwa na Mathieu Roy) (1st ed.). DL_2A - Buluu Publishing. ISBN 9791092789027. [Note: First print edition. Contains selected photographic facsimiles of the original manuscript.]

Mnyampala, M. E. (2013b). *Maisha ni kugharimia* (Kimehaririwa na Mathieu Roy) (1st ed.). Zenodo. https://doi.org/10.5281/zenodo.19738559 [Note: Open-access digital release containing the text and the archival facsimiles.]

Mnyampala, M. E. (2014). *Maisha ni kugharimia* (Kimehaririwa na Mathieu Roy) (2nd ed.). DL_2A - Buluu Publishing. ISBN 9781499199352. [Note: Student edition published without the manuscript photographs to reduce printing costs.]

Mpiranya, F. (2015). *Swahili Grammar and Workbook*. Routledge.

Mugane, J. (2017). The Odyssey of ʿAjamī and the Swahili People. *Islamic Africa*, *8*(1-2), 193-216. https://doi.org/10.1163/21540993-00801005

Mulokozi, M. M., & Sengo, T. S. Y. (1995). *History of Kiswahili Poetry [A.D. 1000-2000]*. Institute of Kiswahili Research, University of Dar es Salaam.

Nabhany, A. S. (2006). Dibaji. In K. W. Wamitila, *Kamusi ya Ushairi* (pp. [3]–[6]). Vide-Muwa Publishers.

Noronha, L. (2009*). Einführung in die Swahili-Literatur. Erster Teil: Prae-Uhuru-Literatur Von den Anfängen bis 1961*. Institut für Afrikawissenschaften der Universität Wien.

Nurse, D., & Hinnebusch, T. J. (1993). *Swahili and Sabaki: A linguistic history*. University of California Press.

Ohly, R. (1985). Literature in Swahili. In B. W. Andrzejewski, S. Piłaszewicz, & W. Tyloch (Eds.), *Literatures in African languages: Theoretical issues and sample surveys*. Cambridge University Press.

Parkar, A. (2021). Arabic-Swahili Hamziyya Manuscripts: Observations on Two Testimonies of the Text. *Manuscript Cultures*, (17), 53-64.

Parry, M. (1971). *The making of Homeric verse: The collected papers of Milman Parry* (A. Parry, Ed.). Clarendon Press.

Poewe, K. (2001), Politically Compromised Scholars, or What German Scholars Working under Missions, National Socialism, and the Marxist-Leninist German Democratic Republic Can Teach Us. *American Anthropologist*, 103: 834-837. https://doi.org/10.1525/aa.2001.103.3.834

Polomé, E. C. (1967). *Swahili language handbook*. Center for Applied Linguistics.

Popper, K. (1959). *The Logic of Scientific Discovery*. Basic Books, Inc.

Pouwels, R. L. (1987). *Horn and Crescent: Cultural Change and Traditional Islam on the East African Coast, 800-1900*. Cambridge University Press.

Pouwels, R. L. (1992). Swahili literature and history in the post-structuralist era: a revisit and riposte to Allen et al. *International Journal of African Historical Studies,* 25(2), 261-283.

Pouwels, R. L. (2002). Eastern Africa and the Indian Ocean to 1800: Reviewing relations in historical perspective. *International Journal of African Historical Studies*, 35(2/3), 385-425.

Pugach, S. (2012). *Africa in translation: A history of colonial linguistics in Germany and beyond, 1814–1945*. University of Michigan Press.

Raia, A. (2020). Prayers uttered down on paper: A stylistic analysis of Swahili devotional lyrics. *Journal for Islamic Studies*, 38(2), 102-140.

Robert, S. (1972). *Sanaa ya ushairi*. Nelson.

Roy, M. (2013a). *KIAMU, archipel de Lamu (Kenya) : Analyse phonétique et morphologique d'un corpus linguistique et poétique*. Presses Académiques Francophones.

Roy, M. (2013b). *Mathias E. Mnyampala (1917-1969) : poésie d'expression swahilie et construction nationale tanzanienne* [PhD Thesis, Institut national des langues et civilisations orientales]. https://theses.hal.science/tel-00778667

Roy, M. (2014). *Mathias E. Mnyampala, 1917-1969, et la construction nationale tanzanienne*. Vol. 1 : Définitions de la poésie d'expression swahilie ; Vol. 2 : Conservatisme et africanisation. DL_2A Buluu Publishing.

Roy, M., & ibn Kadir, U. M. "Mau" A. A. (2013). "*Kiswahili !*" by Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau): Kiamu Phonetic Transcription (IPA) and French Translation. Zenodo. https://doi.org/10.13140/RG.2.1.1598.0400

Roy, M., & ibn Kadir, U. M. "Mau" A. A. (2016). "*Kiswahili !!*": Original Audio Field Recording of the Poem Recited by its Author Sheikh Mahmud Ahmed Abdull Kadir (Ustadh Mau). Zenodo. https://doi.org/10.13140/RG.2.1.3695.1922

Sacleux, C. (1909). *Grammaire des dialectes swahilis*. Procure des PP. du Saint-Esprit.

Sacleux, C. (1939). *Dictionnaire swahili-français*. Institut d'Ethnologie.

Sander, N., & Trenel, I. (1859). *Dictionnaire hébreu-français contenant la nomenclature et la traduction de tous les mots hébreux et chaldéens contenus dans la Bible et dans le rituel des prières journalières*. Bureau des archives israélites.

Shariff, I. N. (1988). *Tungo zetu: Msingi wa mashairi na tungo nyinginezo*. Red Sea Press.

Sonderegger, A. (2011). Racist fantasies: "Africa" in Austrian and German African studies. In W. D. Hund, C. Koller, & M. Zimmermann (Eds.), *Racisms made in Germany* (pp. 123–144). LIT Verlag.

Steere, E. (1870). *Swahili tales, as told by natives of Zanzibar*. Bell & Daldy.

Steere, E. (1894). *A handbook of the Swahili language as spoken at Zanzibar* (A. C. Madan, Ed.). (Original work published 1870)

Thiesen, F. (1982). *A manual of classical Persian prosody: With chapters on Urdu, Karakhanidic and Ottoman prosody*. Harrassowitz.

TUKI (Taasisi ya Uchunguzi wa Kiswahili). (2004). *Kamusi ya Kiswahili Sanifu* (2nd ed.). Oxford University Press.

Vernet, T. (2005). *Les cités-États swahili de l'archipel de Lamu, 1585-1810 : Dynamiques endogènes, dynamiques exogènes* [Doctoral dissertation in Archaeology and Prehistory, in French, Université Paris I - Centre de Recherches Africaines]. HAL Open Science. https://hal.science/tel-01243399

Vierke, C. (2011). *On the Poetics of the Utendi: A Critical Edition of the Nineteenth-century Swahili Poem "Utendi Wa Haudaji" Together with a Stylistic Analysis*. LIT Verlag.

Wallah, W. B. (1988). *Malenga wa Ziwa kuu: Maswali na istilahi za kifasihi*. East African Educational Publishers.

Wamitila, K. W. (2006). *Kamusi ya Ushairi*. Vide-Muwa Publishers Limited.

Weil, G. (1958). *Grundriss und System der altarabischen Metren*. Harrassowitz.

Werner, A. (1917–1920). Swahili poetry. *Bulletin of the School of Oriental Studies, 1*, 113–127.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. Kegan Paul, Trench, Trübner & Co.

Zache, H. (1895). Review: C.G. Büttner, Anthologie aus der Suaheli-Literatur, Berlin 1894. *Zeitschrift für afrikanische und oceanische Sprachen, 1*, 373–375.

Zumthor, P. (1990). *Oral poetry: An introduction* (K. Murphy-Judy, Trans.). University of Minnesota Press. (Originally published in French as *Introduction à la poésie orale*, 1983).

Methodological Note on Computational Assistance

For methodological transparency, a distinction must be drawn between the scientific content of this volume and the computational and editorial tools used in its production.

The intellectual architecture and the empirical research presented here are the original work of the author. To be explicit, the following were conceived by the author: the epistemological and historiographical critique of the discipline; the 3-Layer Theory of Swahili Poetics; the extraction of classical terminology and discourse from the master poets of the Swahili language; the formal syllabic matrices and their algebraic equations, which formalize the metrical principles described by Swahili-language poets themselves — all mathematical formalism in this volume, the matrices, the algebraic equations, and every other notational device used to express them, being likewise the author's own devising; the hierarchy governing the rhyme variables; the formalization of the competing exogenous accentual models; the conditions of falsifiability; the three-step verification protocol; the principle of the prose control corpus; the selection and philological establishment of the corpora; and the logical progression of the arguments. The list is not exhaustive, but it covers what determines the shape of this book.

The empirical core of this volume is the auditory annotation of the corpora. Every syllabification, every accentual reading, every decision to retain or exclude a line, and every confidence mark is the author's own, established by ear on the deposited recordings. No automated procedure was used at any point to determine the auditory reading, and none could have been: the decisions rest on repeated listening and are declared as such throughout.

Following strict academic protocols, the references in this volume were supplied by the author from his own scholarly library, doctoral research, and reading; every citation, and every page range attached to it, was in every case verified by the author against the original publication before use, and no reference was accepted on the authority of a tool.

Large Language Models — principally Anthropic's Claude and Google's Gemini, in the versions available between 2025 and 2026 — were used in four capacities, all strictly under the author's direction and specification.

First, for **statistical computation**. The statistical measures reported in Chapters 9 and 10 — including the distributions of accentual groups, the transition rates, the independence tests, the seeded random draws, and the inter-pass comparisons — were computed using scripts developed with the assistance of Claude. These scripts processed the empirical data — the auditory mappings and syllabifications — exclusively produced and supplied by the author, according to the specifications the author set.

Second, for the **development** of the Python analysis scripts deposited in the accompanying repository. These scripts were also developed with the assistance of Claude and Gemini to implement a measurement protocol defined by the author, processing the data supplied by the author to produce the acoustic figures presented in this volume. The author reviewed the algorithmic logic, modifying and correcting the code where necessary, and functionally validated the scripts to ensure precise control over the specific analytical parameters.

Third, regarding the **English formulation**: English is not the author's first language, and the prose of this volume took shape through a continuous, iterative exchange with these models. The conceptual frameworks, the source material, and the content they express are the author's own; he directed each exchange toward a precise intended wording, revising, rejecting, and requesting alternatives until the result matched it exactly. The models assisted in shaping and refining the linguistic surface of the English prose, while the theoretical choices, the argumentation, and the final conclusions remain exclusively the author's own.

Fourth, as **critical interlocutors**. The models were repeatedly asked to identify weaknesses in the argument and in the verification protocol. Several safeguards of Chapter 9 originated in these exchanges — the rule of non-revision, the deliberately hostile acoustic sample, and the intra-judge control — and were adopted, specified, and implemented by the author.

The author assumes full and final responsibility for the content, arguments, sources, and any remaining errors of this volume.

DL2A Buluu Publishing,
66 avenue des Champs-Élysées,
75008 PARIS,
FRANCE.

ISBN 978-1-291-70670-3
60 EUR
Dépôt légal octobre 2026

Cover picture: Reproduction of a classical Swahili manuscript. Specimen page from the manuscript by Muhammad Kijuma of his version of the *Utendi wa Liyongo* (after 1913). From Harries, L. (1962). *Swahili Poetry*.

Copyright notice / Worldwide Public Domain: This image is a faithful reproduction of a two-dimensional public domain work. In accordance with the threshold of originality required for copyright protection in major jurisdictions, including the United States and the European Union (pursuant to Article 14 of the Directive (EU) 2019/790), this mechanical reproduction does not constitute a new copyrightable work. It remains in the public domain and is not subject to copyright or related rights worldwide.

9 781291 706703

www.dl2a.eu

Has classical Swahili poetry been measured by the wrong ruler for over a century?

For decades, an entrenched lineage of Western philology, spanning from the missionary and colonial theories of Edward Steere and Carl Büttner, through Alice Werner, to contemporary scholars like Clarissa Vierke and Emiliano Minerba, has forced Swahili poetic metrics through the distorted lens of imported accentual systems. *The Procrustean Bed* dismantles this exogenous fallacy. By returning to the material reality of the language and testing these claims against the classical corpus itself, this manifesto for data-driven models shows that the *utenzi*, *wimbo* and *shairi* meters obey the syllabic metrical grids of *vina na mizani*. This is the classical consensus of the poets of the Swahili language themselves, and recovering it frees Swahili poetics from a superimposed academic illusion.



About the Author: As a linguist and a scholar of classical Swahili poetry and manuscript studies, Mathieu Roy conducts research throughout East Africa, combining editorial work on Swahili literature with interdisciplinary methodologies. His expertise spans linguistics, computer science, and the humanities, underpinned by proficiency in Ancient Greek, Arabic, and Swahili. He maintains an active engagement with computational advances. His alias, Mawazoni, is a Swahili term meaning 'in the thoughts' or 'lost in the thoughts'.

DL2A Buluu Publishing
www.dl2a.eu

9 781291 706703