A lexical and phonological reconstruction of Highland East Cushitic including a comparative East Cushitic database

Sosal Altayeb Mohammed Ali, A.

5 Conclusions

This chapter synthesizes the results of Chapters 1–5 and states what they jointly imply for reconstructing Highland East Cushitic. It also separates inherited structures from later changes, where the comparative evidence requires caution. Phonological reconstruction and morphophonological reconstruction are treated here as one analytical system: Segment inventories, prosody, phonotactics, and boundary alternations are interpreted together because the strongest internal diagnostics for reconstruction often come from morphologically conditioned environments.

5.1 What the Study Has Established

5.1.1 Phonological Reconstruction

Chapter 2 reconstructs the phonological system of PHEC on the basis of regular sound correspondences across the five HEC languages (Burji, Hadiyya, Kambaata, Sidaama, and Gedeo). The reconstruction is explicit about degrees of confidence as proto-phonemes are evaluated according to the regularity of the correspondences, the breadth of the supporting examples, and the strength of the comparative evidence.

The reconstruction yields a PHEC inventory with a five-vowel system and contrastive quantity alongside relatively symmetrical consonant subsystems. The consonantal system includes both pulmonic and non-pulmonic series, and the analysis highlights that not all complex correspondence patterns are equally diagnostic. The stops, nasals, and most vowels tend to show the most reliable regularities, whereas sibilant behavior is more complex, interacting with regular sound change, harmony patterns, and (in some cases) areal influence.

Chapter 2 establishes the basic phonotactic and prosodic frame. PHEC accent is reconstructed as weight-sensitive, with later changes in terminal vowels contributing to systematic differences in accent placement across the daughter languages. The result is a coherent picture in which quantity, syllable structure, and prominence are not independent components, but mutually constraining dimensions of the same reconstructed system.

Chapter 3 continues the reconstruction by shifting the evidential focus from lexemes and segmental correspondences to alternations at morphological boundaries. It synthesizes the major morphophonological processes, summarizes their distribution across the HEC languages, and evaluates which

processes are reconstructable to PHEC and which are better treated as later innovations (including changes that may reflect contact).

Two structural domains are especially prominent. First, boundary-driven repairs in stem+suffix combinations frequently create an additional syllable to resolve impermissible consonant sequences, most commonly through vowel or glide epenthesis (including interconsonantal vowel/intervocalic glide epenthesis that must be distinguished from support-verb morphology where relevant). Second, hiatus repair is resolved through glide insertion (***-y-**, with ***-w-** as a conditioned realization near back vowels) or, in a more restricted subset of languages, by glottal stop insertion.

The comparative distribution of these repairs matters. Interconsonantal epenthesis is broadly productive, but the same cognate set does not generally show uniform epenthesis across languages, which limits the reconstructibility of many surface insertions. On the evidence assembled in Chapter 3, only ***-y-** (together with its conditioned realizations) is proposed as a secure PHEC epenthetic segment, while other consonantal insertions are treated as later, language-specific (or contact-induced) changes unless future comparison establishes regular correspondences.

5.1.2 Morphological Reconstruction

Chapter 4 4extends the reconstruction to nominal and verbal morphology. On the basis of reconstructed morphemes and their reflexes across the HEC languages, it argues that the historical change of HEC morphology reflects a complex interaction between inherited morphemes and later changes that may include contact-induced change.

The chapter's primary contribution to the conclusion synthesis is twofold: It consolidates a cross-linguistic inventory of reconstructable morphemes (with explicit gaps and uneven attestations), and it provides a comparative basis for distinguishing what can be projected to PHEC/PNHEC from what is better analyzed as later restructuring. Where the morphological evidence bears on subgrouping or on contact histories, those implications are taken up separately in §§5.3–5.4.

5.2 Concluding Phonology Synthesis

This section synthesizes the phonological and alternations at morpheme boundaries results of Chapters 1–3 as a single system (segmental inventories, prosody, phonotactics, and boundary processes). Chapter 4 supplies the

morphological structures (affixes and paradigmatic environments) that make the boundary alternations in Chapter 3 interpretable as structured evidence rather than isolated irregularities to support the lexical reconstruction. Subgrouping arguments and a discussion on language contact are treated in §5.3.

5.2.1 Phonological System

Chapters 1–3 motivate a simple methodological point with wide consequences for reconstruction: In HEC, the phonological system is most clearly exposed where morphology forces difficult sequences. Segment inventories and prosodic generalizations are not only visible in regular correspondences but also in the ways stems and affixes interact.

The key advantage is interpretive rather than merely descriptive. Boundary alternations show which constraints are active (and stable enough to be reconstructed), and they provide a principled way to diagnose non-initial consonants and consonant clusters that are otherwise obscured by reduction, fusion, metathesis, and variable surface repairs. The morphological analysis in Chapter 4 is crucial here because once the relevant affixes and paradigmatic contexts are identified, seemingly disparate alternations resolve into a small set of recurring strategies.

5.2.2 Segmental Profile of PHEC

At the segmental level, the vowel system is both compact and stable: A five-vowel inventory with a phonemic quantity contrast (treated moraically) is recoverable across the family. The conclusion that matters for later analysis is that vowel quantity and terminal-vowel behavior jointly structure prosody and many surface word forms.

Terminal vowels are a shared property across HEC, but their realization varies by language and morphosyntactic context (case/gender marking) and by diachronic reduction (apocope and devoicing in some cases). This diversity in realizations helps explain why accent targets shift, why word edges are frequent sites of erosion, and why boundary contexts become especially informative.

The reconstructed 21-consonant system is comparably stable in its overall size and typological profile, including both pulmonic obstruents/sonorants and a non-pulmonic series. The analytical lesson from Chapters 1–3 is that consonantal reconstruction is often easiest in positions where the surface is

least stable. Unlike in initial position, the non-initial sounds and clusters are precisely the target where boundary alternations, distributional restrictions, and repair choices provide decisive diagnostics.

5.2.3 Phonotactics and Prosody

Phonotactically, the system can be summarized as a restrictive root template: (C)V(V)(C)(C), with a strong preference for open syllables and a ban on initial clusters. Where clusters occur, they are limited and shaped by sonority sequencing. The broader generalization is that “complexity” is managed rather than eliminated. The HEC family permits certain coda and medial configurations, but it regulates them tightly, which in turn makes repairs predictable at morpheme boundaries.

The prosodic system is best understood as weight-sensitive. Prominence aligns with heavy syllables (those with double vowels or closed ones), but surface placement is frequently restricted to the ultimate/penultimate syllable. The family-wide pattern supports a diachronic scenario in which an earlier “ultimate heavy” target is variably reanalyzed as penultimate prominence when terminal vowels reduce or lose the final consonant needed to bear accent.

5.2.4 Boundary Alternations

Boundary alternations are the meeting point between the abstract phonological system (what sequences are permitted) and the morphological system (what sequences are demanded). Across the languages, the most revealing patterns concentrate in stem + suffix concatenation and in vowel-vowel adjacency (hiatus), precisely where the syllable template and cluster constraints would otherwise be violated.

> Assimilation provides the most direct evidence for active feature pressures at boundaries, and it frequently targets obstruents.
>
> Fusion and gemination show how boundary configurations can be adapted to fit into segmental outcomes (including glottalization in some processes), turning morphological structure into reconstructable phonological evidence.
>
> Palatalization emphasizes the close relationship between segmental modification and the vocalic environment influenced by affixation.

Insertion (epenthesis) and deletion (elision) represent complementary responses to constraint pressure – either add segments to make a form pronounceable or remove material that creates an illicit configuration. However, not every inserted segment is reconstructable; the comparative signal most strongly supports an epenthetic ***-y-*** at the PHEC level, while many other insertion choices appear to be later, language-specific solutions.

Metathesis is common in NHEC and rare in Burji. It is revealing where it clusters in the same prosodic and phonotactic "problem zones", again tying surface diversity to a small set of underlying constraints.

Chapters 1–3 support a conservative principle that treats boundary alternations first as evidence about constraints and the permissible repair processes, and only second as evidence for particular segments.

5.2.5 HEC Sound Classes

The boundary evidence motivates a high-level sound classification that is useful for both description and explanation. The obstruents behave as the main target of feature change, while sonorants behave as the main resource for making forms well-formed.

Obstruents are preferentially affected by feature-changing processes (especially assimilation) and participate readily in boundary outcomes like fusion and gemination. Sonorants, by contrast, appear repeatedly as "phonotactic mediators", so they are favored in epenthesis to satisfy sequencing preferences and to resolve illicit clusters, and they participate in metathesis patterns that improve syllable structure in morphological formations.

Table 9: *Major Sound Classes and Boundary Processes*

Major Class	Minor Class	Assim.	Fusion & gemin.	Metath.	Inser.
Obstruent	Stop	✓	✓		(-ʔ-)
	Non-pulmonic	✓	✓		
	Fricative	✓	✓		
Sonorant	Lateral & Trill	✓		✓	✓
	Nasal	✓		✓	✓
	Glide				✓

Table 10 organizes the consonant inventory by place and manner of articulation within these two major classes. It serves as a compact reference point for the segment sets implicated in assimilation, lenition, and cluster-repair patterns in **Table 9**.

Table 10: *PHEC Consonant Inventory Organized by Class (place × manner)*

		Labial	Alveolar	Post-alveo	Velar	Phary.	Glottal
Obst.	Stop	**b**	**t, d**		**k, g**		**ʔ**
	Non-pulmonic		**t'**	**c'**	**k'**		
	Fricative	**f**	**s, z**	**ʃ**		**ħ**	**h**
Son.	Lateral & Trill		**l, r**				
	Nasal	**m**	**n**				
	Glide	**w**		**y**			

The combined evidence of Chapters 1–3 supports a coherent diachronic picture in which a constrained phonotactic system is maintained through a set of recurrent boundary strategies. The value of the morphophonological record is that it makes those constraints visible and comparable across languages, even where non-initial sounds are unstable on the surface. This synthesis provides the phonological and morphological background needed for the next parts of Chapter 5, where shared innovations are evaluated for subgrouping and where contact-induced patterns are separated from broadly inherited tendencies.

5.3 Subgrouping Evidence

Highland East Cushitic (HEC) is well established as a genetic unit within the East Cushitic family in the comparative literature. As Sasse (1982, p. 9) recalls, "Highland East Cushitic (HEC), a term coined by Bender (1971) to comprise the languages Sidamo, Darasa (Gedeo), Kambata (with two parallel dialects: Timbaro and Alaba), Hadiyya and Burji… grew out of the traditional 'Eastern Sidama group'." In the Language Survey of Ethiopia, "Fleming follows Moreno (1940) in treating HEC as a subgroup in East Cushitic" (Hudson, 1976, p. 246). Hudson (2006, pp. 295–296) further synthesizes the typological profile that characterizes the HEC languages. Here, I advance an innovation-based classification, based on the enhanced coverage afforded by the updated comparative database (§1.3) and detailed reconstruction in the previous chapters.

The subgrouping argument in this study is based on converging evidence across domains rather than on any single diagnostic parameter. The most heavily weighted criteria are shared innovations with restricted distributions, especially when they involve bound morphology or productive alternations (because such patterns are less likely to be independently replicated across languages than isolated lexical similarities).

Accordingly, this chapter distinguishes three kinds of comparative signal:

> **Shared inherited background** (reconstructable to PHEC): processes and distributions that are present across Burji and the NHEC branch and therefore define the diagnostics against which later changes must be evaluated. Measured exceptions supported by Cushitic evidence allow reconstruction to PHEC despite restricted distribution.
>
> **Branch-defining innovations** (at the PNHEC level or below): processes whose distribution aligns with a coherent subgroup and which are absent (or different) in Burji.
>
> **Language-specific changes**: innovations that concentrate in a single language and therefore clarify divergence without themselves defining an internal subgroup.

Because many HEC patterns are morphologically conditioned, the subgrouping logic is explicitly cumulative: phonological innovations (Chapter 2) are interpreted together with morpheme boundary alternations (Chapter 3) and shared morphological markers (Chapter 4), while contact-driven irregularities are treated separately (§5.4).

The lexical evidence is integrated on the same comparative terms. The East Cushitic Database (§1.3) links reconstructions and correspondence judgments to explicit cognate sets and source documentation via stable IDs, so that each claim can be checked against the data upon which it was made. The reconstruction practice also enforces a minimum-attestation threshold: a proto-form is reconstructed only when a cognate set is supported by at least three daughter language cognates within the relevant scope. This has a direct subgrouping consequence. A set attested in Burji plus two NHEC languages can license a PHEC reconstruction without licensing a PNHEC reconstruction; absence at the PNHEC level therefore reflects insufficient internal evidence within the available dataset, not historical absence. A substantial portion of lexical reconstructions are identical at the PHEC and PNHEC levels,

while PNHEC also shows a set of reconstructable lexemes and changes that are not found in Burji.

5.3.1 Phonological Diagnostics

The phonological reconstruction in Chapter 2 establishes that Burji shows the earliest and most regular divergences from the rest of the HEC languages, while the remaining languages share a set of changes that Burji lacks. This asymmetry is the first subgrouping result, supporting an initial split between Burji and a Northern Highland East Cushitic (NHEC) branch.

Chapter 2 also names concrete innovations that help diagnose this split. Three examples that are stated explicitly in the chapter's synthesis are:

> **Shared NHEC innovation:** loss of initial ***k'** and ***ħ**, treated as part of the shared innovation set supporting NHEC unity.
>
> **Burji innovation:** a Burji-specific stop-sibilant metathesis pattern (inherited ***C$_{[velar]}$s** sequence metathesizes in Burji), identified as a distinct innovation of the Burji line after its split from what would become PNHEC.
>
> **Systematic Burji vs. NHEC divergence:** in certain cognate sets, Burji consistently shows **d'** where NHEC languages show **t'**, a recurrent correspondence asymmetry that aligns with the early separation of Burji.

At this level of the synthesis, the critical point is methodological. A subgrouping claim is strongest when multiple independent innovations align with the same split. For this reason, the phonological evidence from Chapter 2 is not treated in isolation here, but taken together with the alternations at morpheme boundaries and morphological innovations.

Chapter 3 provides especially direct subgrouping evidence because morphophonological processes are (a) paradigmatically supported, (b) structurally conditioned, and (c) often distributionally restricted in ways that map neatly onto internal branches.

Table 11: *Morphophonological Subgrouping Diagnostics*

Process/Feature	**Distribution**	**Proto-Level interpretation**	**Subgrouping implication**
Full **-t** assimilation to preceding **C**	Had, Kam, Sid, Ged (NOT Burji)	PNHEC innovation	Supports Burji/NHEC split
t → **c/ʃ**, **k** → **ʃ** palatalization before **i**	Burji only	Burji innovation	Burji-specific
Obstruent + **ʔ** → glottalized geminate (fusion)	Had, Kam, Sid, Ged (NOT Burji)	PNHEC innovation	Supports Burji/NHEC split
Stop + **s** → metathesis	Burji only	Burji innovation	Burji-specific
Sonorant + ʔ → metathesis	Had, Kam, Sid, Ged (NOT Burji)	PNHEC innovation	Supports Burji/NHEC split

Several boundary-driven patterns are shared across Burji and the NHEC branch and can therefore be reconstructed at the PHEC level with relatively high confidence. These include: assimilation of the plurative pronominal suffix to a preceding stem-final sonorant, affrication affecting a nominal derivational suffix, intervocalic lenition of stops, and common repair strategies such as intervocalic/interconsonantal epenthesis (with an epenthetic ***-y-*** established as secure). These shared properties define inherited diagnostics: Similarities in these domains do not, by themselves, constitute evidence for a PNHEC subgroup.

Against that inherited baseline, several morphophonological processes show an exclusive NHEC distribution and thus function as PNHEC diagnostics (or narrower subbranch innovations) that contribute directly to internal classification. In particular, full assimilation of the pronominal suffix ***-t*** after obstruents, the fusion of an obstruent with a following glottal stop yielding glottalized geminates, and metathesis in sonorant + glottal stop sequences are restricted to NHEC and absent in Burji. These innovations jointly support the Burji vs. NHEC split and reinforce the divergence patterns established on phonological grounds.

Burji, conversely, shows morpheme boundary alternations not shared with NHEC. The most notable are the palatalized reflexes before ***i*** (or ***y***) environments and the unique stop + sibilant metathesis pattern mentioned above. These patterns confirm that Burji diverged before the NHEC innovations took place.

5.3.2 Morphological Diagnostics

Morphological innovations are among the strongest types of evidence for subgrouping because shared markers and paradigmatic structures are less likely to be replicated independently across languages. Chapter 4 demonstrates that morphology is also a locus for contact-induced change. In this synthesis, Chapter 5 therefore distinguishes (a) inherited morphology that defines the PHEC baseline from (b) branch-specific innovations and (c) contact-driven lookalikes.

5.3.2.1 Nominal Morphology

In nominal morphology, Chapter 4 treats gender consonants as among the most secure reconstructions. PHEC ***-k** (masculine) and ***-t** (feminine) are reconstructable across demonstratives and definiteness markers, and are attested in gender and case morphology. Their comparative stability and inherited background provide a stable baseline against which later developments can be evaluated rather than serving as a diagnostic innovation by themselves.

The singulative and deverbative ***-ʃ** is also widely attested throughout HEC with diverse allomorphy. Therefore, it is a secure PHEC reconstruction. The only distinction between the singulative and the deverbative is in their templates: nominal stem + ***-ʃ** → singulative, while verbalized stem + ***-ʃ** → deverbative.

The remaining exponents distribute within smaller groups comprising 2–3 languages: Sidaama, Gedeo, Burji (collective ***-ey-e**), Hadiyya, Sidaama, Gedeo (plurative ***-uw-a**). However, their reconstruction to PHEC still qualifies as reasonable based on wider Cushitic evidence. The same applies to the stem-final reduplication plurative (***C_i~C_i-a**) found in all HEC and other Cushitic languages, but not in Burji.

The genitive exponent ***-ʔ** is reconstructed to PHEC with arguable certainty, with clearer segmental evidence in some languages and indirect reflexes in others (e.g., glottalization or accentual effects after glottal loss).

5.3.2.2 Pronominal Morphology

The pronominal paradigm provisionally reconstructed in **Table 8** (§4.3) reveals three defining structural properties. First, suppletion is confined to first and second persons (distinct roots (**an-*, **at-*, **naʔ-*, **ati-*)), contrasting with

the third person built on ***is-**. Second, case marking strategies differ by subsystem: The nominative employs vowel suffixes (***-i, *-u**), the accusative uses suppletive forms (**ke, *is-o, *is-e*) or vowels, while the possessive relies on the genitive glottal (**ʔ-y*, **ni-ʔ-*, **kin-ʔ-*). Third, metathesis affects possessive glottal stems in the PHEC → PNHEC transition (e.g., PHEC **kin-ʔ-* becomes PNHEC **kiʔn-* in 2pl), serving as a diagnostic PNHEC innovation. The 3pl possessive **isn-ʔ-* resists metathesis due to consonant clustering, triggering vowel insertion across the reflexes.

These morphological patterns reveal that the pronominal paradigm participated fully in systematic sound changes, differentiating PHEC from PNHEC (Ch. 3). Metathesis of possessive glottals represents paradigm-specific conditioning (glottal + nasal sequences), while 3pl cluster resolution demonstrates phonotactic repair. The interplay between phonological regularity (the metathesis rule applies) and morphological analogy (paradigm leveling in 2pl PHEC **ati-n-*, PNHEC **aʔ-n-*) provides evidence for the integrated morphophonological reconstruction undertaken in this study.

5.3.2.3 Reflexive Morphology

Reflexive morphology is diagnostic for subgrouping because Section 4.2.3 explicitly separates an inherited exponent from later, distributionally restricted developments. A core reflexive marker ***-ʔ** is securely reconstructable to PHEC on the basis of broad distribution and regular morpheme boundary alternations.

Within NHEC, §4.2.3 treats the reflexive alternations at morpeheme boundaries as part of a regular conditioning profile, i.e., metathesis after sonorant-final stems and fusion after obstruent-final stems (yielding glottalized geminate outcomes). In addition, complex reflexive formations such as **-ak'k'** / **-it't'** are described as restricted to Kambaata and Hadiyya and are therefore candidates for narrower internal NHEC innovations (cf. §4.2.5.9).

With regard to the contact discussion in §5.4, see also §4.2.3.4, which reports reflexive variants involving **-(V)d'** in Burji, Gedeo, and Sidaama, and treats them as borrowed from Oromo and not reconstructable to PHEC or PNHEC. These forms should be excluded from subgrouping arguments and treated instead as a contact signal.

5.3.2.4 Verbal Derivation

The verbal subsections (§§4.2.4–4.2.5) emphasize a recurring diachronic asymmetry: Change often targets individual exponents while preserving derivational templates (slot structure). This supports reconstructing the verbal template architecture to PHEC, even where particular extensions are unevenly distributed. Section 4.2.5 also invokes a grammaticalization trajectory (support verb → verbalizer → valency marker), which provides a coherent explanation for why template structure can remain stable while particular formations are replaced or reanalyzed.

For subgrouping, the template architecture represents the PHEC baseline rather than diagnostic innovations. Diagnostic potential lies in restricted lower-level innovations, in particular exponents or allomorphs such as **-it't'** in Kambaata and Hadiyya.

Table 12: *Morphological Subgrouping Diagnostics*

Exponent/structure	**Distribution**	**Proto-level interpretation**
Passive ***-am**	All HEC	PHEC (secure)
Causative ***-s**	All HEC	PHEC (secure)
Causative **-ʃ**	Kam, Sid, Ged	Innovation (narrow)
Gender ***-k** (m), ***-t** (f)	All HEC	PHEC (secure)
Genitive ***-ʔ**	Had + HEC (indirect)	PHEC (arguable)
Plurative ***C_l~C_l-a**	NHEC	PHEC (reasonable)
Plurative ***-uw-a**	Had, Sid, Ged	PHEC (reasonable)
Collective ***-ey-e**	Sid, Ged, Burji	PHEC (reasonable)
Singulative / Deverbative ***-ʃ**	All HEC	PHEC (secure)
Reflexive *-ʔ	All HEC	PHEC (secure)
Reflexives **-ak'k' / -it't'**	Had, Kam	Innovation (narrow)
Reflexive **-Vd'**	Burji, Ged, Sid	Contact-induced
Verbal templates (exponent turnover)	All HEC	PHEC (reasonable)

5.3.3 Subgrouping and Revised EC Labels

Chapters 1–4 establish an initial split between Burji and an NHEC branch comprising Hadiyya, Kambaata, Sidaama, and Gedeo. The evidence is cumulative: Phonological divergence patterns are reinforced by multiple NHEC-specific morphophonological innovations and by Burji-specific developments that concentrate outside the NHEC distribution.

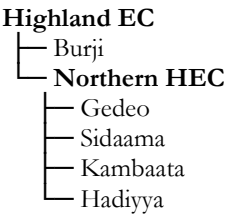

Figure 2: Highland East Cushitic Family Tree (introduced in §1.1.3)

More fine-grained internal subgrouping within NHEC requires careful filtering for contact effects (§5.4) and a fuller integration of the morphological distributions. For this reason, the present synthesis treats the Burji vs. NHEC split as the primary subgrouping result, leaving more detailed internal branching to a subsequent analytical step. This revises and refines Hudson's (2007, 1989, 1981) HEC group relationships, where Hadiyya-Kambaata and Sidaama-Gedeo share common nodes.

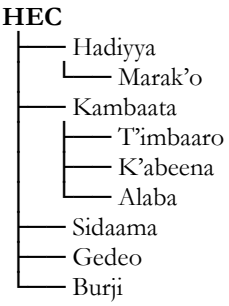

Figure 5: *Hudson's Family Vine Tree (2007, 1989, 1981)*

Hudson's (1981) "family vine" envisions the HEC languages as a gradient continuum or chain. In the family vine, Burji is the closest to Gedeo, Gedeo to Sidaama, Sidaama to Kambaata, and Kambaata to Hadiyya based on a quantitative count of shared grammatical features from a comparative survey (Hudson, 1976). This vine-like arrangement, illustrated via relative strengths

of shared traits (e.g., phonological, morphological, lexical), emphasizes dialect geography and wave-like diffusion over discrete branching, while acknowledging the challenges of distinguishing innovations from retentions or borrowings. Thus, Hudson's quantitative enumeration of features treats all features equally without prioritizing innovations. The current study refines this by focusing on diagnostics, confidence evaluation, and explicit contact diagnostics, while excluding probable areal diffusions such as **-(V)d'** in Burji, Sidaama, and Gedeo. By layering contact filtering, the current study aligns the vine's gradient insights with a genealogical hierarchy, where PNHEC emerges as a robust node.

Lexically, Hadiyya-Kambaata and Sidaama-Gedeo pairs share the most cognates (confirming Hudson's subgrouping), followed by Hadiyya-Kambaata-Sidaama. However, no solid shared innovations support internal subgrouping of the four NHEC languages: Hadiyya, Kambaata, Sidaama, and Gedeo. Contact signatures provided false signals for many of Hudson's shared features used in his subgrouping. NHEC-internal contact likely contributed to the shared lexicon in these groups, while the signals Burji and Gedeo share derive from their common donor, Oromo.

In sum, HEC forms a secure genealogical unit defined by a cross-domain bundle of innovations (Ch. 2–4). Internally, the primary split separates Burji from PNHEC, supported by a number of innovation bundles and, independently, by loan distributions. Within PNHEC, there is no signal strong enough to allow the positing of a reliable subgrouping based on shared innovations. Hadiyya and Kambaata preserve relatively conservative profiles that aid the reconstruction.

I use revised traditional subgroup labels in this study for four EC subgroups. Where traditional labels use a single language name as a cover term for a subgroup node, I adopt descriptive, geographically anchored terms for clarity (**Northern EC** for Saho-Afar; **Central EC** for Oromoid; **Sagan** for Konsoid, **Southern EC** for Yaaku-Dullay). I also propose these terms as a practical alternative for future comparative work, since they make the intended scope of each node more transparent while avoiding unintended asymmetries by remaining strictly terminological. Other than the new "Northern HEC" designation, these labels function as descriptive terms and do not imply new subgrouping claims.

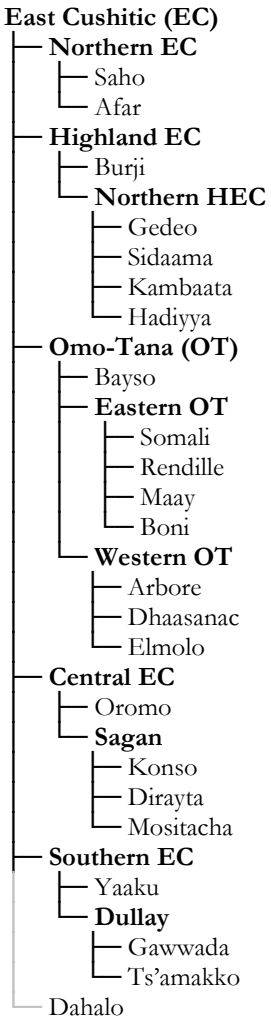

Figure 6: *Revised East Cushitic Subgroups Labels*

This study does not posit any arguments against Lowland East Cushitic, and even proposes an arguable diagnosis for the Lowland East Cushitic node based on the non-pulmonic evidence discussed in §2.1.2 pending further research.

The next section (5.4) turns to the contact ecology and the layered borrowing history (involving Oromo, Amharic, and Omotic) that shaped the lexicon and left recognizable donor signatures without overturning the genealogical signal.

5.4 Language Contact and Diffusion

Language contact is treated in this study as a pervasive historical factor rather than as an occasional anomaly. Interaction with neighboring languages, especially Oromo and Omotic, has left detectable traces across HEC, though the degree and region of influence vary by language and by domain.

On general comparative grounds, the most expected contact areas are (a) the lexicon (loanwords and loan translations), (b) distributional clustering of variants by geography rather than by inheritance, and (c) irregular reflexes or double outcomes that resist a regular correspondence explanation.

The most direct contact evidence is in the lexicon, where borrowed forms enter without exhibiting inherited correspondences and introduce segments or clusters that are otherwise marginal. For this reason, lexical distributions are interpreted cautiously when used as evidence of subgrouping. The discussion in §5.3, therefore, treats lexical signals as strongest when they align with innovation-based evidence from phonology and morphology.

Debates about Burji's position also underscore why contact must be modeled explicitly in subgrouping. Sasse (1982, pp. 9–10) frames the relationship of Burji to HEC as a long-standing classificatory problem and notes that Burji shows, in some respects, stronger affinities to Lowland East Cushitic than other languages classified as HEC. Hudson (1981, p. 116) likewise treats language contact as a potential contributor to Burji's divergence, while emphasizing that Burji's grammatical divergence is not straightforwardly attributable to borrowing. In this study, such observations motivate contact filtering and conservative interpretation of lexical distributions rather than any relaxation of innovation-based subgrouping criteria.

Regarding contact, the literature has repeatedly noted substantial Oromo influence on parts of the HEC lexicon, especially in Burji and Gedeo. Hudson (1989, p. 8) describes all five HEC languages as showing apparent borrowings from Oromo and Amharic, with Burji and Gedeo standing out for their especially strong Oromo lexical influence. Hudson (1976, p. 276) similarly observes: "In Burji and Darasa (more so in Darasa) [Gedeo], the lexicon is considerably penetrated by Oromo loans. But for neither language should these loanwords inspire doubts about its genetic affiliation". This study also confirms that Burji's and Gedeo's divergence is due to heavy contact with a (Lowland) East Cushitic language, and the evidence comes from Oromo. Contact does not necessarily appear as a broad disruption of correspondence

structure. It more often surfaces as restricted irregularity, geographically clustered variants, or outcomes that sit outside the inherited system and therefore require a donor hypothesis. This matters for interpretation because the same surface irregularity can reflect borrowing, diffusion, or the coexistence of inherited reflexes beside borrowed ones.

Morphology is integrated into this discussion for two reasons. First, borrowed segments (e.g., **d'**) can be amplified through morphological analogy, increasing their apparent regularity. Second, where morphological restructuring aligns with known contact zones, it must be evaluated cautiously before being treated as evidence for subgrouping. Chapter 4 provides explicit examples of this logic in the reflexive domain, where **-(V)d'** variants are treated as Oromo-induced developments and are excluded from proto-level reconstruction.

The working diagnostic for separating contact from inheritance follows the evidential logic stated in §1.5. In practice, the analysis privileges: (a) regular correspondences within HEC to establish inherited structure; (b) distributional irregularities as potential indicators of contact; (c) wider East Cushitic comparison to establish what counts as regular before positing a contact source; and (d) areal comparison where it clarifies direction and plausibility of borrowing. For example, comparing the non-pulmonic PHEC ***t'** and PHEC ***c'** from Chapter 2 (§2.1.2), their similarities include (a) being proto-phonemes reconstructed as inherited, (b) showing regular correspondences across HEC, (c) showing position-independent retention (initials/non-initials), (d) having complications from contact (Oromo, Amharic), and (e) having correspondences in the rest of EC. The key differences, on the other hand, are summarized as follows:

Feature	PHEC *t’	PHEC *c’
Primary status	Inherited non-pulmonic	Inherited post-alveolar non-pulmonic
Contact complications	Irregular **d’** and **c’** reflexes in subset	Amharic loanwords with regular **c’**
Oromo influence	Causes irregular **d’** and **c’** in Burji, Gedeo, Sidaama, and “double reflexes”	Causes “double reflexes”
EC correspondences	Strong, widespread	More parallels, mostly regular but less systematic data
Non-Cushitic parallels	More parallels	Fewer parallels
Distribution	Appears more common	Appears less common (fewer examples cited)

The crucial difference is the irregular vs. regular **c’** reflexes. The irregular **c’** (from ***t’**) is contact-induced, limited to Burji, Gedeo, and Sidaama, and found on specific lexemes. In contrast, the regular **c’** (from ***c’**) is retained, found in all five HEC languages, and has regular correspondences. This distinction justifies reconstructing both ***t’** and ***c’** as separate proto-phonemes.

Where a pattern is fully regular within HEC and supported by substantial comparative evidence, the analysis generally does not attempt to partition the outcome into “inheritance” versus “contact” components, because any ancient influence may have been fully regularized. In such cases, parsimony favors analysis as inherited unless external evidence provides a stronger reason to posit contact. Where irregularity is present, however, contact remains the preferred hypothesis and is incorporated into confidence evaluation rather than excluded by default.

5.5 Methodological Implications

Across the reconstruction (Ch. 2–4), claims are not treated as equally secure by default. The analysis follows the probability-based evaluation framework established in §1.5 and assigns confidence evaluations to reconstructions based on (a) the regularity of the correspondence patterns, (b) the number and distribution of supporting examples, and (c) the strength of the comparative evidence.

In practice, this means that the study prioritizes convergence of evidence from multiple sources: regular correspondence sets, distributional constraints, and morpheme boundary alternations that expose underlying structure. When correspondence patterns are complicated by harmony, positional

conditioning, or areal convergence, the evaluation framework makes that uncertainty explicit rather than forcing a categorical reconstruction.

Although contact and inheritance are treated as a continuum in the history of HEC (rather than binary alternatives), reconstructions are assessed with an attention to distinguishing inherited structure from later innovation. This chapter maintains this separation: §5.2 synthesizes the phonological and morphophonological system, while §§5.3–5.4 develop the subgrouping and contact interpretations.

Boundary alternations provide a different kind of comparative evidence than lexical correspondences do. They are valuable precisely because they are systematic, productive, and tightly conditioned by morphological concatenation. In the HEC data, stem + suffix environments repeatedly trigger repairs that would be hard to motivate from citation forms alone, especially when the surface outcomes differ across languages.

Two points are vital for reconstruction. First, boundary-driven syllable creation (often via vowel/glide material) and stem-final consonant alternations clarify what counts as an impermissible cluster in each system and show which repair strategies are productive. Second, hiatus and alternation-repair identify a limited set of recurrent solutions whose distribution is itself comparative evidence, e.g., glide insertion patterns (***-y-**, with ***-w-** as a conditioned realization near back vowels) contrast with more restricted glottal stop insertion. Boundary alternations also force the analysis to separate proto-level structure from later changes. Some starred input patterns used in summaries (§3.5) are best treated as schematic or intermediate stages rather than direct proto-forms, because subsequent assimilation and other boundary processes may obscure the underlying composition.

A recurring limitation in the reconstruction is that the absence of an alternation is rarely decisive evidence for absence at the proto-stage. Attestation gaps, uneven descriptions, and differences in analysis all distort the surface distribution of boundary processes. For this reason, negative evidence is treated cautiously throughout the study.

Reconstruction, therefore, proceeds by ranking hypotheses rather than maximizing proto-inventory. When a process is broadly productive but does not align across cognate sets, it is more plausibly a later innovation (including contact-induced developments) than a retained proto-level feature. This

conservative stance keeps the reconstruction consistent with the comparative requirements stated in §1.5.

5.6 Contributions and Recommendations

This study presents a lexical, phonological, morphophonological, and morphological reconstruction of Highland East Cushitic grounded in an explicit evaluation framework. The phonological component reconstructs a PHEC segmental inventory and prosodic profile on the basis of regular correspondences across the HEC languages. The morphophonological component shows how morpheme boundary alternations provide evidence for constraints and repair strategies that are not always recoverable from citation forms. Chapter 4 supplies the paradigmatic environments that make many alternations interpretable and keeps phonological generalizations grounded in attested morphological structure rather than in isolated word forms.

A central research output is the East Cushitic Lexical Database (§1.3), developed as part of this dissertation. The database is publicly available as an Excel database and an interactive web application, and it links dissertation IDs to individual lexical entries so that reconstructions, cognate sets, and supporting sources can be traced directly. Its scope is explicitly comparative across East Cushitic: It includes 23 East Cushitic languages, 356 PHEC and 375 PNHEC lexical reconstructions, in addition to reconstructed PWR parallels, and brings together over 8,500 compiled lexical entries from more than 100 scholarly sources and limited amounts of (my own) data collected (between 2024–2025) from speakers of Dahalo, Arbore, and Boni. Methodologically, it encodes reconstruction status as working hypotheses and makes uncertainty explicit by assigning each reconstructed form a Recon.Eval. confidence score (0–4), tied to correspondence regularity, semantic uniformity, and complications such as irregular correspondences or limited attestation. The EC and PWR parallels are only proposed as potential cognates.

One limitation concerns semantic and functional reconstruction. The present study is primarily a reconstruction of forms, correspondences, and distributional patterns – it hardly attempts a systematic reconstruction of semantic change criteria or a comparative reconstruction of the full range of functions associated with individual morphological exponents. These domains are nevertheless important for strengthening reconstructions and for distinguishing inherited structure from later reanalysis. In practice, semantic uncertainty is kept visible rather than normalized away: The evaluation

framework takes into account semantic stability, and glossing in examples is kept conservative (simplified glosses are used where items have multiple meanings, with the full range of attested glosses available in the database; database IDs in square brackets link examples to entries). A fuller semantics- and function-centered reconstruction would require richer and more uniform documentation than is presently available across the comparative sources assembled here.

A further limitation is uneven evidence. Some processes are widely productive but are not consistently realized across cognate sets, which restricts how far particular surface segments (especially epenthetic consonants) can be projected to the proto-level. Similarly, gaps in attestation and differences in descriptive coverage mean that negative evidence cannot be treated as decisive.

Several broader limitations also follow from the infancy of HEC comparative studies and from the design choices required by reconstruction. For some domains, the available evidence supports multiple competing hypotheses; adopting one hypothesis is sometimes necessary for consistency with dependent analyses, even where the comparative signal remains incomplete. In addition, the present study's bottom-up focus on HEC necessarily limits the degree to which wider East Cushitic comparison can be used to judge marginal correspondence patterns. Where wider comparison is likely to be decisive, the conclusions are framed conservatively.

In addition to addressing the above limitations, future work can strengthen the reconstruction in three ways. First, expanding and normalizing comparable datasets across languages would make it easier to test whether marginal correspondences reflect systematic conditioning or later remodeling. Second, deeper comparison beyond HEC, where relevant, can help distinguish inherited features from later areal convergence, especially in domains where contact is pervasive. Third, more detailed work on paradigms and morphosyntactic conditioning of terminal vowels and accent would refine the diachronic link between word-edge erosion and prosodic reorganization.

5.7 Conclusion

This study shows that reconstructing the Highland East Cushitic lexicon and phonology is most successful when the comparative method is paired with careful attention to paradigms and boundary alternations. The resulting system is typologically coherent: a constrained phonotactic profile maintained

by a limited set of repair processes, a prosodic pattern tied to syllable weight and sensitive to terminal-vowel change, and a set of morpheme boundary alternation processes that expose structure where surface forms are least stable.

The comparative database published with this research provides a resource for Cushitic linguistics and a foundation for further inquiry into Afroasiatic historical linguistics, particularly by making comparative evidence and reconstruction evaluation explicit and traceable through stable identifiers.