



Universiteit
Leiden
The Netherlands

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

Sosal Altayeb Mohammed Ali, A.

Citation

Sosal Altayeb Mohammed Ali, A. (2026, July 9). *A lexical and phonological reconstruction of Highland East Cushitic including a comparative East Cushitic database*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/4307764>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4307764>

Note: To cite this publication please use the final published version (if applicable).

2 Phonology

This chapter reconstructs the phonological system of Proto-Highland East Cushitic (PHEC). To achieve this goal, the sound correspondences across five HEC languages (Burji, Hadiyya, Kambaata, Sidaama, and Gedeo) are analyzed by investigating the regularity of sound change. The reconstruction applies the probability-based evaluation framework established in §1.5. The confidence evaluation terms are used accordingly for each proto-phoneme based on the regularity of correspondences, the number of supporting examples, and hence the strength of comparative evidence.

Because reconstructed proto-phonemes (marked with *) are used from the outset to formulate correspondence hypotheses, they should be treated here as working analytical labels. The full reconstructed phoneme inventory is presented in §2.5.1 (**Table 3**) and §2.5.3 (**Table 5**), and a consolidated class-based summary of the consonants is provided in Chapter 5 (**Table 10**).

This chapter demonstrates how the historical-comparative method reveals both inherited and contact-induced changes in the HEC phonological systems, as outlined in the methodological foundations in Chapter 1. The analysis follows an integrated approach that recognizes language contact as a pervasive factor shaping phonological change. Such an approach avoids the previous pitfalls of treating contact and inheritance as binary alternatives. Most of the sound correspondences support straightforward inheritance from PHEC. However, others are explained as a result of prolonged contact with neighboring languages, particularly Oromo. The reconstruction distinguishes between these scenarios by examining comparative evidence from the broader East Cushitic family where needed (see §2.1.2).

This chapter reconstructs the PHEC consonants, vowels, syllable structure, and accent. Section 2.1 establishes the consonant inventory of PHEC, examining pulmonic and non-pulmonic consonants across labial, alveolar, post-alveolar (palatal), velar, and glottal positions. Section 2.2 addresses the vowel system, including single, double, and terminal vowels. The two shorter sections that follow deal with the syllable structure (Section 2.3) and accent of PHEC (Section 2.4), respectively. The final Section 2.5 provides summaries and concluding remarks on the reconstructed phonology. The reconstruction evidence is presented systematically: the proposed proto-form, regular reflexes in the daughter languages, supporting examples from the database (with unique IDs in brackets), and a discussion and evaluation of correspondence regularity and reconstructed lexemes.

The terminological distinctions established in §1.5 are maintained throughout the reconstruction. Database entries are referenced with unique IDs in brackets (e.g., PHEC [72] **lam-* ‘two’), enabling readers to examine full comparative sets in the database. Otherwise, the data source is cited. For further examples, only reconstructed lexemes are provided; readers can check the full context in the database by following the respective ID. Cross-references to other chapters (Ch. 3 for alternations at morpheme boundaries and Ch. 4 for morphology) appear where an apparently phonological question can be answered by morphology.

2.1 HEC Consonants

This section includes five subsections organized by manner or place of articulation:⁴ pulmonic alveolars and post-alveolar (§2.1.1), non-pulmonic alveolars and post-alveolar (§2.1.2), labials (§2.1.3), glides (§2.1.4), and velars, glottals, and the pharyngeal (§2.1.5). The order is methodological rather than proportional. Subsections are arranged for analytical progression, and cross-references are used where related evidence and interpretation are developed in different sections. These subsections establish the consonant inventory of PHEC through the analysis of correspondence patterns. Each proto-consonant reconstruction is examined based on regularity, attestation, and comparative evidence, and evaluated using the probability-based framework (§1.5.1).

To establish the synchronic foundation for the diachronic reconstruction that follows, a brief overview of the consonant inventories attested across the modern HEC languages is provided in **Table 1**. Transcription symbols follow the conventions presented in §1.5.3.

The HEC languages collectively have 27 consonants, with individual languages ranging from 23 to 26 consonants.

⁴ I have organized these subsections by manner and place to reflect patterns in the distribution of reflexes.

Table 1: *Highland East Cushitic Consonant Inventory*

	Labial	Alveolar	Post-alveolar	Velar	Glottal
Stop	p, b	t, d		k, g	ʔ
Non-pulmonic	p', b'	t', d'	c'	k'	
Fricative	f	s, z	ʃ		h
Affricate			c, j		
Lateral & Trill		l, r			
Nasal	m	n	ɲ		
Glide	w		y		

This study analyzes 24 of these consonants because Sidaama, Burji, and Kambaata **ɲ** and Burji **p, b'** are unattested in the database (§1.3). Several consonants require particular attention in reconstruction: the distribution of non-pulmonic consonants across languages, the variation in sibilant inventories (**ʃ** presence/absence), and the restricted attestation of certain phonemes (**ɲ, p, b'**). These synchronic patterns guide the identification of which consonant features are inherited versus innovative in each language.

As outlined in the chapter introduction, the reconstruction applies an integrated approach to contact and inheritance. This approach is particularly relevant for non-pulmonic consonants, where Oromo and Omotic influence is evident (§2.1.2). The reconstruction proceeds through comparative evidence from within HEC and, where relevant, broader East Cushitic. Proto-forms are documented with database IDs (e.g., PHEC [48] **tum-* 'to pound'), to enable cross-referencing to full comparative evidence.

2.1.1 Pulmonic Alveolars and Post-alveolar

This section analyzes the regular sound correspondences of Burji, Hadiyya, Kambaata, Sidaama, and Gedeo to reconstruct the pulmonic alveolar and post-alveolar consonants in PHEC. While the stops **t* and **d* are securely reconstructed with consistent reflexes, the fricatives present more complex patterns. The voiced fricative **z* has merged with **d* in most languages, and the voiceless alveolar fricatives **s* and **ʃ* exhibit sporadic retention alongside merger patterns. Confidence evaluations (secure, reasonable, arguable, or tentative) for these reconstructions follow the framework outlined in §1.5.1. This section examines stops (§2.1.1.1), fricatives (§2.1.1.2), laterals and trills (§2.1.1.3), and nasals (§2.1.1.4), establishing secure reconstructions for **t*, **d*, **l*, **r*, **s*, and **n*, while the fricatives **z* and **ʃ* present more complex patterns requiring careful evaluation. Additionally, the phonological processes

outlined in the phonological processes chapter (Ch. 3), such as assimilation and palatalization, are essential for interpreting the reflexes of the common suffixes consisting of alveolars and post-alveolar.

2.1.1.1 Stops

The alveolar stops ***t** and ***d** reconstruct to PHEC, with regular correspondences across all five daughter languages.⁵

PHEC ***t**

The reconstruction of the voiceless alveolar/dental stop ***t** in PHEC is supported by regular reflexes across the HEC languages. Reflexes of ***t** appear in both initial and non-initial positions, as illustrated by the following cognate sets:

	[48]		[521]		[605]	
PHEC	*tum-	'to pound'	*it-	'to eat'	*galat-	'thanks'
Burji	tum-	'to pound'	it-	'to eat'	galataa	'thanks'
PNHEC	*tum-	'to pound'	*it-	'to eat'	*galat-	'thanks'
Ged	tum-	'to pound'	it-	'to eat'	galata	'thanks'
Sid	tum-	'hit, strike'	it-	'to eat'	galata	'thanks'
Had	tum-	'to pound'	it-	'to eat'	(galata) ⁶	'thanks'
Kam	tum-	'to pound'	it-	'to eat'	galata	'thanks'

These reflexes illustrate the stability of ***t** across HEC languages. However, alterations involving **t** at morpheme boundaries reveal processes such as assimilation, gemination, and, particularly in Burji, palatalization – see §3.1 for alternation processes affecting alveolars.

Further examples supporting the reconstruction of PHEC ***t** include: PHEC **[40]** **tuf-* 'to spit', **[130]** **tuul-* 'to pile up', **[71]** **k'utañ-* 'thorn', **[87]** **k'ot-* 'to dig', **[180]** **utub-* 'pole, center-pole of house', **[142]** **tam-n-* 'ten', and PNHEC **[254]** **tun~tumm-* (← **tum~tum-*) 'fist', PHEC **[462]** **meet-* 'scraping board for use in preparing ensete', and PNHEC **[460]** **beet-* 'son, child'.

Overall, the reflexes of PHEC ***t** across HEC languages demonstrate its stability as a phoneme in both initial and non-initial positions, and it can be

⁵ According to Sasse (1982, p. 16), Burji **t** and **d** are dental [**t̪**] and [**d̪**] rather than alveolars. This might be an areal pronunciation reported for Rendille, Dahalo, and coastal Bantu languages as well (see Nurse, 1985).

⁶ The parentheses indicate a problematic parallel (§1.3.3 and §2.1.1.3), and the Hadiyya reflex **l** is irregular in this example.

reconstructed as a secure proto-phoneme. However, Burji's distinct palatalization (§3.1.2.1) introduces an innovative layer to this phonological picture.

PHEC *d

Parallel to *t, the voiced alveolar stop *d is also securely reconstructed for PHEC. The reflexes are consistent in initial and non-initial positions.

	[672]		[325]		[378]	
PHEC	*dul-	'to skin/ slaughter'	*duud-	'mute (person)'	*dan-	'be able'
Burji	dul-	'pierce'	duuda	'dumb'	dan-	'be able'
PNHEC	*dul-	'to skin/ slaughter'	*duud-	'mute (person)'	*dan-t-	'be able'
Had	dur-	'slaughter, kill'	duuda	'mute'	-	-
Kam	dul-	'slaughter'	duuda	'mute'	dand-	'be able'
Sid	dul-	'skin, flay'	duuda	'mute'	danda?	'be able'
Ged	dul-	'skin'	duuda	'blunt, dull'	dand-e?-	'be able'

The above reflexes reveal a retention of *d, particularly in the initial position. In the case of [378] *dan- 'be able', the absence of the final d in Burji compared to the PNHEC form *dan-t- is due to a fossilized denominative suffix -t, which underwent voicing assimilation as discussed in §3.1.1.1. The verbalizer -? in Sidaama and Gedeo (§4.2.5.1), and potentially the reflexive in Burji dan-?- 'able for oneself' (§4.2.3.1), indicates later independent suffixation.⁷

Further examples supporting the reconstruction of PHEC *d include PHEC [316] *dam- 'branch, stick', [114] *debe?- 'to belch', and [206] *awd- 'threshing floor', PNHEC [389] *duk- 'to carry, to creep', [629] *bad- 'to mistreat', [606] *gubeed- 'thigh', [399] *dawam- 'red', and [436] *ad- 'aunt'.

The reflexes of *d in HEC underscore its stability across HEC. For the relevant alternations involving some non-initial reflexes of *d, see §3.3.1.2.

2.1.1.2 Fricatives

The reconstruction of fricatives in PHEC presents a more complex picture than the stops, particularly with regard to the voiceless alveolar fricatives. The

⁷ see Gawaada *tana?*- and Oromo *dandaya* 'be able', probably with denominatives as well (§3.1.1.1).

fricative system poses methodological challenges: distinguishing merger from retention, evaluating contact effects, and assessing the reliability of isolated cognates. The analysis below prioritizes systematic correspondences over sporadic attestations.

PHEC *z

The reconstruction of *z rests on regular correspondences in Kambaata and Alaaba, contrasting with the *z → d merger in Burji, Hadiyya, Sidaama, and Gedeo. The following comparanda illustrate both the retention pattern and the contrast with d reflexes:

	[118]		[14]		[139]	
PHEC	*zaak-	'to swim'	*az-	'milk'	*wozan-	'heart'
Burji	daah-	'swim'	ada	'milk'	wodana	'heart'
PNHEC	*zaaz-,	'to swim,	*az-	'milk'	*wozan-	'heart'
	*zaak	flow'				
Had	(daad-)	'spill, flow'	ado	'milk'	wodano	'heart'
Kam	(zaaz-)	'spill, flow'	azu-ta	'milk'	wozana	'heart'
Sid	daak-i	'swim'	?ado	'milk'	wodana	'heart'
Ged	daak-	'swim'	ado	'milk'	(onna)	'heart'

In [118], the crucial evidence for *z is Alaaba **zak-** 'swim', which simultaneously preserves the *z reflex (parallel to Kambaata) and the 'swim' meaning (parallel to Burji/Sidaama/Gedeo). The Kambaata and Hadiyya stem-final consonant in **zaaz-** / **daad-** 'spill, flow' may reflect assimilation following the lenition of **k** to **h** (§3.2.2). Semantically, a different Sidaama lexeme, **daad-** 'spill, flow' (cf. Gedeo **dan-am-** 'spill, flow'), aligns better with the Hadiyya and Kambaata meaning, unless borrowed from Hadiyya/Kambaata.⁸ Without Alaaba, the case for [118] would be weak; with it, the *z reflex and the 'swim' gloss are linked in a single witness, and the strongest support for PHEC *z then comes from [14] **az-** 'milk' and [139] ***wozan-** 'heart', where the correspondences are straightforward. In [139], Hadiyya shows variation between **wodano** and **wadano** 'heart', while Gedeo **onna**, without the initial **w**, is likely borrowed from Oromo **onn-ee** 'heart'.

Despite the limited retention of *z in only Kambaata and Alaaba, both closely related languages within NHEC, the broader EC comparanda (e.g., Arbore, Saho **z**, and PHEC *z correspondences) support the reasonable

⁸ Oromo **daaka** 'swim' provides a further indication of the possibility that the Burji, Gedeo, and Sidaama parallels are borrowed.

reconstruction of ***z** to PHEC, with a subsequent merger in most daughter languages.

The retention of ***z** in Kambaata and the related language provides crucial evidence for reconstructing it for PHEC. However, the widespread merger with **d** in Burji, Hadiyya, Sidaama, and Gedeo requires an explanation. Two competing analyses have been proposed to account for the distribution. Hudson (1989, p. 8) proposes reconstructing ***dz** for PHEC, arguing that it simplifies the derivation: ***dz** → **z** in Kambaata/Alaaba and ***dz** → **d** elsewhere. Sasse's (1976, 1979) simpler ***z** reconstruction is preferred here. Hudson's derivation is also parsimonious, but ***z** accounts for the Kambaata/Alaaba reflex directly, i.e., no deaffrication step is required and is consistent with the absence of affricates elsewhere in the reconstructed PHEC inventory. Hudson's ***dz**, by contrast, introduces a proto-segment with no other consonant classification support.

Further examples of PHEC ***z** reflexes include: PHEC [438] **zuug-* 'to scrape, tan (hide)', [511] **zal~zal-* 'trade, business', [581] **zar-* 'to tear', [330] **sooz-* 'dawn', [383] **orz-* 'hair of animals', and [444] **bagaz-* 'spear'.

HEC languages other than Kambaata and Alaaba do have instances of **z**, but only in loanwords. For example, Anbessa Teferra (2000, p. 13) notes that **z** in Sidaama occurs only in Amharic loans, and Hudson (1976, p. 248) similarly observes its rare attestation in Hadiyya and Gedeo except in borrowed words. Bender (2020, p. 116) also notes that **z** is found rarely in Burji and is restricted to loans.

The comparative evidence supports a reconstruction of PHEC ***z** with limited retention and widespread merger with **d**. This aligns with a broader pattern in East Cushitic languages, where only a handful of languages retain a reflex of **z** against the majority that show a corresponding **d**.

PHEC *s

The voiceless alveolar fricative ***s** reconstructs securely to PHEC based on consistent reflexes across all five languages in both initial and non-initial positions.

38 A Lexical and Phonological Reconstruction of Highland East Cushitic

PHEC	[5] *san-	'nose'	[117] *saʔ-	'cow'	[7] *is-u	'he' (3sg.m.nom)
Burji	suna	'nose'	saa	'cow'	ʔisi	'he, it'
PNHEC	*san-	'nose'	*saʔ-	'cow'	*is-u	'he' (3sg.m.nom)
Had	sane	'nose'	saáyya	'cow'	isse	'he' (honorific)
Kam	sanu-ta	'nose'	saʔa	'cow'	ís/ísu	'he'
Sid	sano	'nose'	saa	'cow'	ʔisi	'he'
Ged	sano	'nose'	saayya	'cow'	isi	'he'

Further examples include: PHEC [723] **sek-* → **seh-* (→ **seh-s* (caus)) 'adorn, decorate', [20] **sazeñ-* 'three', [39] **suk-ʔ-* 'to spin', [637] **sim-* 'ensete tree 1st stage', [330] **sooz-* 'dawn', [108] **sunk-* 'to kiss', [638] **sor-* 'red soil', [350] **sang-* 'ox, bull' (see Amharic *sānga*, Oromo *sangaa* 'ox for beef'), [461] **has-* 'to want, seek', [633] **has y-s-* 'be necessary', PNHEC [601] **hos-* 'to pass/spend the day', PHEC [603] **boos-* 'fireplace', and PNHEC [33] **osal-ʔ-* 'to laugh'.

The reconstruction of **s* for PHEC is secure, supported by consistent correspondences across all five HEC languages.

Sporadic cases of **s* → *f* show Oromo influence on Burji because it is a typical Oromo sound change. Such an influence is attested in two HEC numerals, though these words with *f* are not attested in Oromo. The shift is not phonologically conditioned and cannot be predicted from the environment. It is item-specific, contact-induced replacement reflecting the broader Oromo influence on Burji rather than a Burji-internal sound change. PHEC *s* otherwise remains *s* in Burji.

PHEC [20] **sazeñ-* 'three' → Burji *fadiya*, PNHEC **sazeñ-* 'three' → Had *saso*, Kam *sasu*, Ged and Sid *sase*.

PHEC [467] **ñons-* 'nine' → Burji *wonf-a*; PNHEC **ñons-* 'nine' → Had *honso*, Kam *honsu*, Sid *honse*.

There is one probable instance of hypercorrection reverse change **f* → *s* in Burji:

PHEC [120] **fil-ʔ-* 'to comb' → Burji *fil-*, *sil-*; PNHEC **fil-ʔ-* 'to comb' → Had *fit'-*, Sid *fit't'-*, Kam *fiil-*, Ged *fil-*.

Sasse (1982, p. 166) argues that Burji *sil-ʔ-* 'to comb' is unrelated to Oromo *fil-* 'to comb', which has original *f* from PEC **fil-* 'separate'; his account implies

Burji preserves only the sibilant form. However, Getahun Addo (2017, p. 307) documents Burji *fil-* ‘to comb’ with the expected regular **f** reflex of PHEC and PEC, coexisting with *sil-ʔ-* in the same language. The presence of both forms in Burji is what undermines Sasse’s separation and supports the hypercorrection reading taken here: *fil-* is the inherited form, and *sil-ʔ-* is a Burji-internal hypercorrection back to ***s** under sustained Oromo contact.

PHEC ***ʃ**

The evidence for reconstructing PHEC ***ʃ** is more tenuous than that for other fricatives, paralleling the situation in the rest of East Cushitic (see Sasse, 1979, p. 31). Sporadic initial and non-initial correspondences are provisionally taken as evidence for traces of retained and re-established PHEC ***ʃ** via innovated and borrowed lexicon with no clear East Cushitic cognates (see §2.1.5.1 (PHEC ***k'**)).

The analysis below progresses through these dimensions: first establishing the basic correspondence sets that justify the reconstructions, then addressing the distributional and irregular patterns and their explanations, followed by systematic treatment of contact phenomena, areal effects, and finally geminate attestations.

	[8]		[497]		[249]	
PHEC	*ʃool- ⁹	‘four’	*ʃakol- , *ʃol-	‘be easy, light’	-	-
Burji	foola	‘four’	(jakk-ad’-)	‘be easy’	-	-
PNHEC	*ʃool-	‘four’	*ʃoll ʔ	‘be easy’	*ʃik’ y / ʔ	‘to approach’
Had	sooro	‘four’	ʃoll-eʔ-	‘be easy’	ʃik’aaʔ-	‘approach’
Kam	ʃoolu	‘four’	ʃalal-	‘be easy’	ʃik’ y	‘budge’
Sid	ʃoole	‘four’	ʃot-	‘be light’	ʃik’k’ifi	‘approach’
Ged	ʃoole	‘four’	ʃolled’-	‘be light’	ʃik’k’a	‘approach’

In PHEC [8] ***ʃool-** ‘four’, Oromo influence in the form of ***s** → **f** is evident in Burji *foola*, despite Oromo having [1035] *afur* ‘four’. Burji **f** is not a direct reflex of PHEC ***ʃ**; it traces via the chain **ʃ** → **s** → **f**, where the second step is the Oromo-influenced **s** → **f** shift already established under PHEC ***s** above. The Burji **s** ~ **ʃ** alternation documented across these sets is what allows the intermediate stage to remain visible; the **f** output is the contact-induced terminus of the same chain. Hadiyya’s **s** reflex is unique in HEC, so it might

⁹ PHEC [8] ***ʃool-** ‘four’ via an intermediate ***ʃohl-** if related to PD ***salaḥ** ‘four’ both from earlier ***salḥ**? Vowel quality is problematic.

be an irregular change.¹⁰ Examples of the verbalizers like ʔ and the other reflexes of the denominative in [497] are discussed in §4.2.5.¹¹ Further examples of initial ʃ include: PHEC [95] *ʃomb-* ‘lung’ and PHEC [237] **ʃirb* ‘to dance and sing’.

The reconstruction of *ʃ for PHEC is arguable, given the complex merger patterns, its sporadic retention, and the areal influences documented below.

All arguable retentions of PHEC *ʃ in the present dataset are word-initial. In intervocalic (non-initial) position, there is no direct reflex contrast between inherited *s and *ʃ in HEC, which points to an extensiveness of the merger in this environment.

PHEC [252] **hasaaw-* (← **hafaaw-*) ‘speak, talk’ → Burji *hasaaw-* ‘to talk’; PNHEC **hasaaw-* ‘speak, talk’ → Had *hasaaw-* ‘chat’, Kam *haasaaww-* ‘discuss’, Sid *hasaaw-* ‘speak, chat, talk’, Ged *haasoʔ* ‘speak, chat’.

Set [252] is therefore treated as an indirect case: the non-initial s can be interpreted as the outcome of earlier *ʃ only by comparison with set [229] PHEC **haf~haf-* ‘to whisper’. The reconstruction **hafaaw-* is inferential, not based on direct retention of a non-initial *ʃ contrast, and any semantic link between the two sets is treated as tentative. An additional comparative indication comes from Oromo: since PEC *s regularly yields Oromo f (Sasse, 1979, p. 33), the s in Oromo cognate *haasa(a)y-*, *aasaay-* ‘speak, chat’ and *(h)asaasa* ‘to whisper’ cannot derive from PEC *s and is better traced to PEC *ʃ, which independently supports the reconstruction of (Pre-PHEC) **hafaaw-*. Burji s (instead of expected ʃ) thus reflects leveling/contact pressure from the same Oromo form rather than direct inheritance of *s.

For some instances of non-initial ʃ adjacent to consonants, it is difficult to determine if they are part of the root or suffixed allomorphs of the causative (see §3.1.2.3) or causative denominative (§4.2.5.7) because they are almost exclusively found on verbs. An example is the likely metathesized suffix -s/-ʃ in Burji [45] *fiisk-* and Hadiyya *fifkaʔa* (intr) ‘to whistle’. The reflexes of ʃ

¹⁰ Similar s ~ ʃ variation in NHEC appears in the Amharic loanword [394] *ʃənkurt* ‘onion’ → Burji *ʃunkurtee*, Ged *ʃunkurte*, Sid *sunkurta*, Had *ʃunkurutta* / *ʃinkurutta* / *sunkurutta*, Kam *sunkutta-ta* / *ʃunkurta-ta*. Hadiyya and Kambaata show free variation between ʃ ~ s, of the Amharic ʃ, while Sidaama shows only s.

¹¹ On the notation of bare y and ʔ (e.g., **ʃik’ y* / ʔ ‘to approach’) as support verbs and not as suffixes, see §1.5.

adjacent to consonants can be approximately generalized as NHEC **ʃ**, and Burji **s** or **ʃ**.

PNHEC [693] **darf*- ‘to swell’ → Had *daff*- ‘swell’, Sid, Ged *darf*-.

PHEC [229] **haf~haf*- ‘to whisper’ → Burji *hasaas*- ‘whisper’; PNHEC **haf~haf*- ‘to whisper’ → Had *hafafa wac*-, Kam *haaaff*-, Sid *haaaf*-, Ged *hafaf*-.

PHEC [109] **f*- ‘to kill’ → Burji *siy*-; PNHEC **f*- ‘to kill’ → Had, Kam *f*-, Sid *ʃi*, Ged *ʃi-ya*.

Again, the Burji word is closer to Oromo (*h*)*asaasa* ‘whisper’ in [229], which might explain the **s** reflex. The full reduplication is reconstructed provisionally to account for the double V₂ reflexes attested in Burji, Kambaata, and Sidaama (§2.2.2.2), which would otherwise be unmotivated under a partially reduplicated proto-form, once the weak non-initial **h** and the absence of **h** in clusters are taken into account. This does not deny the onomatopoeic character of the root but treats it as having stabilized into a regular reduplicated stem at the proto-level, which is why it remains a usable witness here.

The EC parallels of PHEC [109] **f*- ‘to kill’ show **ʃ** in Konso *ikaʃ*-, *iff*- ‘kill’, **f** in Saho *igdife* ‘to kill’, and **s** in Oromo *ajeess*-, *ijees*- ‘kill’, Rendille *agiis* ‘to kill’, and Elmolo *ékis* ‘to kill, extinguish’. Sasse (1982, p. 167) reconstructed this root as a prefix verb in PEC **-gʃi* ‘to kill’. If PEC had **-gʃi*, that means PHEC ***ʃ** became Burji **s** via ***g** devoicing to ***kʃ** and subsequent lenition to ***hʃ**, followed by the loss of the fricative **h** (§2.1.5.2). Burji *siy*- and Gedeo *ʃi-ya* may be compounds including the support verb **y* ‘to say, do, be’. Kambaata also has *ʃiyye* (imp.pl) ‘kill!’.

Proto-East Cushitic ***ʃ** and ***s** regularly became Oromo **s** and **f** and PNHEC ***ʃ** and ***s**, respectively. The NHEC reflexes are sporadic at times, so Oromo reflexes are used to determine some proto-phonemes, such as the following:

Oro *sama* ‘to rot, mould’, PHEC [523] **fam*- ‘to rot, to be wet’ → Burji *sam*- (from Oro), *fam*-; PNHEC **fam*- ‘to rot, be wet’ → Had, Kam, Ged, and Sid *fam*-.

Oro *saalf-ad’d’a* ‘be shy, be ashamed’, PHEC [450] **ʃaal*- ‘be shy, be ashamed’ → Burji *ʃa(a)ʔ*- ‘be shy’, *saalfad*- ‘be embarrassed’;

42 A Lexical and Phonological Reconstruction of Highland East Cushitic

PNHEC **saal-* ‘be shy, be ashamed’ → Kam *saal-*, Sid *saal-*, *saalfat-*, Ged *saalfat-*.

Oro *finc’aan* ‘urinate’, PHEC [170] **sum ʔ* ‘urinate’ → Burji *sinʔ-aaw-*, PNHEC **fum ʔ* ‘urinate’ → Had *fum-eʔ-*, Kam *fum-aʔ-*, Sid *fumi*, *finc’i*, Ged *jiʔn-*.

Oro *reefa* ‘corpse’, PHEC [329] **reñ-f-* ‘corpse’ (see [184] **reñ-* ‘die’) → Burji *reesi*, *reefa* ‘corpse’; PNHEC **reñ-f-* ‘corpse’ → Had *leefa*, *leeffa* ‘corpse’, Kam *reʃa*, Sid *reeffa*, Ged *reenʃa*.

PHEC [450] **faal-* ‘be shy, be ashamed’ is mainly supported by Burji and Oromo reflexes, and so is reconstructed assuming that they are strictly regular. Burji, Sidaama, and Gedeo *saalf-* ‘be shy, be ashamed’ is from or influenced by Oromo *saalf-ad’d’a* ‘be shy, be ashamed’ (§2.1.1.2).¹² However, PHEC [170] **sum ʔ* ‘urinate’ is challenging because Burji shows irregular reflexes, with **ʃ** found on the noun *finʔaa* ‘urine’ (see [145]) and **s** on the verb *sinʔ-aaw-* ‘urinate’ (see Amharic *fänna* ‘urinate’). Getahun Addo (2017, p. 631) documented Burji *sinʔa* ‘urine’ with a regular **s**. The noun [329] **reñ-f-* ‘corpse’ is derived via the derivation suffix **-f** (§ 4.1.1) and challenges the full reliance on Oromo since it is an example of skewed **f** reflex via **ʃ** and **s**, not a direct ***s** (see Ehret, 1987, pp. 139–141).

The following discussion is not an account of any single set already presented; it covers a transversal **s** ~ **ʃ** pattern that crosscuts the inheritance and contact analyses of §2.1.1.2 and connects to the broader obstruent-harmony hypothesis developed in fn. 36. Hayward’s (1988, 1997) sibilant-harmony framework, introduced in Hayward (1988, pp. 274ff) and described in Hayward (1997, pp. 115–118) as an Omotic areal feature adopted by early HEC speakers, provides an arguable account of certain irregularities. In many Omotic languages, sibilant obstruents organize into “palatal” and “non-palatal” sets, maintaining strict segregation within roots. The change of ***t’** → **c’** in Burji *c’eeji* (← PHEC [18] **t’iig-* ‘blood’) likely reflects harmony triggered by the palatalized stem-final ***j** (← ***g**) (§3.2.1, fn. 36). The analogous obstruent-harmony hypothesis of fn. 36 is extended beyond the place of articulation to voicing, irregularities such as PHEC [20] **sazeñ-* ‘three’ → PNHEC **saseñ-* ‘three’ → Had *saso*, Kam *sasu*, Ged and Sid *sase* explicable

¹² A further challenge occurs when one looks at the variants with both **s** and **ʃ** in Burji *saalfad’-*, *faalfad’-* ‘to shy, be feeling fear, become timid’ (Getahun Addo, 2017, p. 735).

as a voicing-harmony effect. This extension goes beyond classic sibilant harmony and is offered tentatively.

The above discussion of PHEC ***ʃ** covers its distribution, inheritance, and contact. Initial **ʃ** has largely merged with **s** except in remnant cases, e.g., PHEC [95] *ʃomb*- ‘lung’, words without cognates in the rest of EC (except Oromo), e.g., PHEC [523] **ʃam*- ‘to rot, to be wet’, PNHEC [249] **ʃik*’*y*, **ʃik*’*ʔ* ‘to approach’, PHEC [639] **ʃaf*-, **ʃaf-ʃaf*-*ʔ* ‘to shake’, and loanwords, e.g., Amharic *ʃəmbəra* ‘chick-peas’, *ʃənkurt* ‘onion’. There is no intervocalic **ʃ** since it merged with **s**. Exceptions include loanwords such as Amharic *käsäl* ‘charcoal’ → PNHEC [704] **kaʃil*- ‘embers, charcoal’. The irregular **s** and **ʃ** correspondences reflect multiple interacting factors along the contact/inheritance continuum. Inherited distinctions from PHEC underwent both sporadic changes (sibilant harmony, an areal Omotic feature adopted by early HEC speakers) and lexicon-specific influences (borrowing from and contact with Oromo, where PEC ***ʃ** → Oromo **s**).

Finally, for both PHEC ***s** and ***ʃ**, there are no reconstructable geminate consonants, except **ass*- ‘to do’, which is geminated due to an old fusion with a preceding consonant via derivation from PHEC [425] *ʔ* ‘to do’ + **-s**. Synchronically, geminate **ss** is found as a reflex of originally single ***s**, at least in the third-person pronouns in Hadiyya and Kambaata, while variation between **ʃʃ** and **ss** is observed in Gedeo and Sidaama in two or three loanwords.

2.1.1.3 Laterals and Trills

The consonant inventory in **Table 1** shows both the lateral **l** and the trill **r**. Their reconstruction to PHEC is secure, supported by consistent retention across all HEC languages except for Hadiyya, which exhibits a unique set of sound changes affecting these sounds.

i. Lateral and trill rules in Hadiyya

- (a) **l** → **r** / **V** **_** **V**
- (b) **r** → **l** / **#** **_**
- (c) **rr** → **ll** (non-initially)

These rules explain the observed reflexes in Hadiyya: **l** → **r** intervocalically, **r** → **l** initially, and geminate **rr** → **ll** non-initially. These Hadiyya-specific changes distinguish it from the other four HEC languages, which preserve the

44 A Lexical and Phonological Reconstruction of Highland East Cushitic

***l**, ***r** contrast consistently. The changes represent independent innovations within Hadiyya rather than shared NHEC innovations, next to the retention of distinct ***l** and ***r** in Kambaata, Sidaama, and Gedeo cognates in all positions.

PHEC ***l**

Here are examples of initial and non-initial single PHEC ***l** and its reflexes across HEC:

	[209]		[137]		[705]	
PHEC	*laaf-	'soft, weak'	*lal-	'cattle'	*kul-	'to tell'
Burji	laafa	'soft, weak'	lál-i	'cattle'	kulʔud	'to stutter'
PNHEC	*laaf-	'soft, weak'	*lal-	'cattle'	*kul-	'to tell'
Had	laafa	'weary, infirm'	laro	'cattle'	kur-	'tell'
Kam	laafa	'soft, weak'	lalu	'cattle'	kul-	'tell'
Sid	laafa	'weak'	laló	'cattle'	kul-	'tell, confess'
Ged	laafa	'soft, weak'	lalo	'cattle'	kul-	'tell'

In Hadiyya, examples **[137]** and **[705]** show a non-initial reflex **r** of PHEC ***l** in accordance with the rule i(a): ***l** → Hadiyya **r** / **V_V**.

Additional supporting evidence for PHEC ***l**: PHEC **[51]** ***lokk-**, ***lekk-** 'foot, leg', **[72]** ***lam-** 'two', **[38]** ***liñ-** 'six', **[4]** ***gulib-**, ***gilub-** 'knee', PNHEC **[33]** ***osal-ʔ-** 'to laugh', and HEC **[136]** ***eel-** 'pond, pool'.

Geminated ***ll** is retained in Hadiyya.¹³

PHEC **[1]** ***ill-** 'eye' → Burji *illa*, *ila*; PNHEC ***ill-** 'eye' → Had *ille*, Kam *illi-ta*, Sid *ʔille*, Ged *ille*.

PHEC **[150]** ***baall-** 'feather' → Burji *baallee*; PNHEC ***baall-** 'feather' → Had *baalleʔe*, Kam *baalli-ta*, Sid *baallicco*, *balle*.

PHEC **[276]** ***ball-** 'brother-in-law, father-in-law' → Burji *bolla*; PNHEC ***ball-** 'brother-in-law, father-in-law' → Had, Kam, Sid, Ged *ballo*.

¹³ Cognates of **[1]** and **[150]** in the rest of EC have a non-geminated **l**, except Oromo *baallee*, *baalʔee* 'feather' and Dhaasanac *baalle*, *baal* 'feather'. The geminated HEC **ll** in **[1]** is plausibly the reflex of final **l** reduplication. A morphological alternative involving a fossilized suffix **-t** is set aside, as PHEC **[175]** ***kalt-** 'axe, hoe' would be a counterexample.

Other examples of geminated ***ll** include PNHEC [191] **ull-* ‘earth’, PHEC [382] **oll-* ‘neighbor’, PHEC [173] **billaa-t-* ‘butterfly’, and PHEC [202] **tall-* ‘only, alone’.

***l** is retained in ***lC** clusters in Hadiyya.

PHEC [175] **kalt-* ‘axe, hoe’ → Burji *kaltee* ‘axe’; Had *kalceʔe* ‘certain part of the plow’, Kam *kalta* ‘hoe’.

PNHEC [298] **el ʔ* ‘to be useful’ → Had *eʔl-* ‘be better, improve’, Sid *el-* ‘to grant, make donation; use, be useful to’, Ged *el-* ‘be better, improve’.

Had [604] *c’alc’al*, Kam *c’alc’al y-* ‘to hang’.

In loanwords, non-initial ***l** is borrowed as **l** in Hadiyya.

PHEC [349] **gabal-* ‘border, limit’ → Burji *gabalaa*; PNHEC **gabal-* ‘border, limit’ → Had, Kam, Sid, Ged *gabala*.

PHEC [605] **galat-* ‘thanks’ → Burji *galataa*; PNHEC **galat-* ‘thanks’ → Had, Kam, Sid, Ged *galata*.

The source of [349] is unknown, while [605] is probably from Oromo *galata* ‘thanks’ given the irregular Hadiyya reflex **l**.¹⁴

Overall, PHEC ***l** is retained in initial position, in clusters, and when geminate in all of HEC. It is only changed in intervocalic position in Hadiyya. Thus, it is a secure reconstruction.

PHEC ***r**

Here are examples of initial and non-initial PHEC ***r** and its reflexes across HEC:

¹⁴ One might wonder if it was **galt-* hence lack of ***l** → **r** in Hadiyya? Or was it transferred from Kambaata into Hadiyya?

46 A Lexical and Phonological Reconstruction of Highland East Cushitic

	[184]		[431]		[110]	
PHEC	*reñ ?	'to die'	*raʔ-	'ripen, be ripe'	*barr-	'day'
Burji	rey-, re-	'to die'	raʔ-	'be ripe'	bari	'dawn'
PNHEC	*reñ-	'to die'	*raʔ-	'ripen, be ripe'	*barr-	'day'
Had	leh-	'to die'	liʔ-is-	'become ripe'	balla	'day'
Kam	reh-	'to die'	reʔ-	'be ready (food)'	barraa	'day'
Sid	re-, reʔ-	'to die'	raʔ-	'ripe'	barra	'day'
Ged	rey-, re(ʔ)-	'to die'	raʔ-	'ripen, be ripe'	barra	'day'

Rule i(b) states that initial ***r** becomes **l** in Hadiyya, as the above examples demonstrate.¹⁵

Further examples of non-initial ***r** include: PHEC [189] ***k'araar-** 'be bitter', PNHEC [42] ***k'eer-** 'long, tall', PNHEC [127] ***kontor-** 'to snore', PHEC [122] ***gor-** 'thorn bush', PHEC [141] ***k'ar-** 'to sharpen', PNHEC [212] ***ar-** 'husband', PNHEC [223] ***eger** 'to wait', PNHEC [224] ***t'ur-** 'dirt of body/clothing', PNHEC [345] ***raar-** 'shout, scream', and PNHEC [65] ***faras-**, ***faras-f-** (sgv) 'horse'.

In a consonant sequence ***rC**, ***r** does not change in Hadiyya.

PHEC [85] ***arg-** 'to lend/borrow' → Burji *arg-*, *erg-* 'send'; PNHEC ***arg-** 'to send' → Had *hereg-* 'to share/to lend', *arguug-* 'to chase', Kam *arg-*, 'lend/borrow something that can be returned', *arguug-* 'chase away', Sid *arg-* 'lend/borrowed (other than money)', Ged *erg-* 'send'.

PHEC [161] ***bark-** 'head-support, pillow (of wood)' → Burji *bórk-e*; PNHEC ***bark-** 'head-support, pillow (of wood)' → Had *barkuma*, Kam *borkaanu*, Sid *barko*, Ged *barko* (cf. Oro *bark-oo* 'headrest').

PHEC [220] ***c'ark'-** 'rags, cloth' → Burji *c'ark'aa* 'rag, worn-out cloth'; PNHEC ***c'ark'-** 'rags, cloth' → Had *c'ark'a* 'cloth, rag', Kam *c'ark'a* 'piece of cloth, shred of cloth', Ged *c'ark'e* 'rags, cloth'.

Hadiyya [85] *hereg-* 'to share/to lend' involves reduplication and insertion of initial **h-** and second vowel **e**. The Hadiyya forms are treated as an inherited

¹⁵ Each of the above examples has one or more related PHEC roots (see the appendices). For set [431], the Hadiyya and Kambaata cognates show a potential vowel ablaut (§2.2.2), and the Sidaama variant *leʔ-* 'ripe' has an irregular initial **l**. Additionally, in PHEC [610] ***roob-** 'hippopotamus' → Burji *roobee*; PNHEC ***roob-** 'hippopotamus' → Had and Kam *looba*, Sid *rooʔe*, Ged *roop'e*, the Kambaata reflex **l** is irregular and possibly due to Hadiyya influence.

reflex of PHEC *arg-* with internal reshaping (reduplication and **h** prosthesis), not as borrowings. The Oromo cognate *erga* 'send' has influenced the semantics and supplied the vowel reflexes in the Gedeo and Burji variants. A similar influence might have been exerted on the vowel reflex (**a** ~ **o**) of the widespread root in the second example [161]. That is in addition to the potential irregular rounding conditioned by the preceding labial. Parallel to PHEC [220] **c'ark-* 'rags, cloth', Amharic has *c'ärk* 'cloth' and Oromo has *c'ark'ii* 'ragged, cloth'.

Geminated **rr* becomes Hadiyya **ll** in accordance with rule i(c) **rr* → **ll** non-initially as seen in PNHEC [110] **barr-* 'day' → Hadiyya *balla* 'day, date, time' above and the following examples:

PNHEC [777] **dirr-* 'to descend, go down' → Had *dill-*, Kam, Sid, Ged *dirr-*.

PHEC [228] **harr-* 'donkey' → Burji *harree*; PNHEC **harr-*, **harr-i-ff-* 'donkey' → Had *halla*, Kam *harri-ta*, Sid *harre*, Ged *hare*.

PHEC [9] **arrab-* 'tongue' → Burji *arraaba*; PNHEC **arrab-* 'tongue' → Had *allabo*, Kam *arrabi-ta*, Sid *ʔarrabo*, Ged *arrabo* (see PHEC [548] **arraab-* 'to lick').

PNHEC [190] **arr-* 'sun' → Had *ellin-co*, Kam *arru* 'sunlight', Sid, Ged *ariffo*.

PNHEC [226] **c'aac'urr-*, **c'iic'oorr-* 'chick(en)' → Sid *c'aac'c'urre*, Had *c'iic'oolla*, *c'iic'oora*, Kam *c'iic'oora*.

Based on the cognates and rule i(c), Hadiyya [226] *c'iic'oolla* is the regular variant, while *c'iic'oora* with a single **r** has no parallel. The above examples show that PHEC **r* changes to **l** (rule i(b))¹⁶ and **ll** (rule i(c)) in Hadiyya in initial position and when geminate, respectively. Otherwise, it is retained across all HEC languages, supporting its secure reconstruction to PHEC.

In conclusion, the changes of the lateral and trill are confined to Hadiyya and can be summarized as follows:

¹⁶ There are two instances of free variation between non-initial **l** and **r** in Hadiyya: PNHEC [397] **kikkil-* 'armpit' → Had *kikkiro*, *kilkilo* and PHEC [121] **gol-* 'wall' → Had *gorteʔe*, *goratanna*, *goleta* 'wall'.

		initial	non-initial	_C
*l	→	l	r	l
*r	→	l	r	r
*ll	→	-	ll	-
*rr	→	-	ll	-

Single ***l** and ***r** merged with **l** word-initially and **r** non-initially (intervocally), and geminate ***ll** and ***rr** merged with **ll** (non-initially). Thus, synchronically, Hadiyya has no initial **r** (Tadesse Sibamo, 2015, p. 22) except in a few loanwords (see Ritter, 2007, p. 111).¹⁷ The contrast between the old single ***l** and ***r** is only preserved before some heterogeneous consonants, i.e., in ***lC** and ***rC** clusters in Hadiyya.

2.1.1.4 The Nasal

PHEC ***n**

The reconstruction of PHEC ***n** is secure despite its defective distribution in initial position. While initial ***n** is rare in native HEC vocabulary, occurring primarily in pronominal forms, e.g., [188] **ni-ʔ-* ‘our’ and suffixes, the consonant shows consistent reflexes across all five languages in non-initial positions

	[188]		[2]		[54]	
PHEC	*ni-ʔ-	‘our’ (pl.poss)	*an-i	‘l’	*min-	‘house’
Burji	ninka/ninku, ninci/ninta	‘our’	ʔáni	‘l’	mina	‘house’
PNHEC	*ni-ʔ-	‘our’ (pl.poss)	*ani	‘l’	*min-	‘house’
Had	ni	‘our’	ʔáni	‘l’	mine	‘house’
Kam	nii, -nne	‘our’	án	‘l’	mini	‘house’
Sid	ninké, -nke	‘our’	ʔani	‘l’	mine	‘house’
Ged	noʔoki, noʔot’t’i	‘our’	ani	‘l’	mine	‘house’

Moreover, broader East Cushitic comparanda demonstrate regular correspondences for ***n** (see PEC [54] **min-* ‘house’), confirming its reconstructibility to PHEC. The restricted distribution in initial position likely reflects a lexical gap rather than a phonotactic constraint, as loanwords freely allow initial **n**, e.g., Oromo [152] *naacca* ‘crocodile’ in HEC.¹⁸ The nasal ***n**

¹⁷ The evidence for reconstructed initial ***r** is limited in the available data.

¹⁸ It is also found on the isolated roots PNHEC [1032] **nigga* ‘tendon’ and PHEC [1033] **naac’a* ‘joke’.

also plays a productive role in HEC morphology, e.g., nominal marking §4.1 (see also genitive §4.1.4.1 and epenthetic **-n-** §3.4.2), further supporting its secure reconstruction.

Other cognate sets show: PHEC [745] **-n* ‘with’, PHEC [66] **anuun-* ‘breast’, PHEC [67] **ang-* ‘hand, arm’, PHEC [5] **san-* ‘nose’, and PHEC [628] **ann-* ‘father’.

2.1.1.5 Summary

The reconstruction of alveolar and post-alveolar consonants in PHEC provides insights into the phonological history of the Highland East Cushitic languages. While the stops **t** and **d** exhibit clear and consistent reflexes, the fricatives, particularly **z** and **ʃ**, present a more complex picture of merger, retention, and potential borrowing. The systematic sound changes affecting lateral **l** and trill **r** in Hadiyya highlight the independent phonological innovations within that language. The nasal **n** has a limited distribution word-initially, yet remains stable in other positions and even plays a significant role in morphology. The following section (2.1.2) examines the non-pulmonic alveolars and post-alveolar, where contact effects and analogical developments require particularly careful analysis.

2.1.2 Non-pulmonic Alveolars and Post-alveolar

Having established the pulmonic consonants, this section examines the non-pulmonic alveolar and post-alveolar consonants, where the relationship between inheritance and contact is more complex. The existing literature describes them as glottalized, glottalic, ejective, retroflex, and implosive consonants, but here I use the term “non-pulmonic” consonants (including the velar **kʼ**). This reserves “glottalized” for secondary non-pulmonics derived through morpheme boundary alternation processes (detailed in §1.5.2). This section begins with a synthesis of key discussions in the existing literature on the non-pulmonic consonants, then presents an analysis of the historical development of these consonants in the HEC languages.

Previous scholars have attempted to reconstruct non-pulmonic consonants across various Cushitic and Omotic proto-languages. The following list synthesizes those key reconstructions:

50 A Lexical and Phonological Reconstruction of Highland East Cushitic

Proto-Language	Non-pulmonics	Source
Proto-Highland EC	*d', *t', *c', *k'	Hudson (1989)
Proto-Lowland EC	*D, *q	Black (1974, p. 280)
Proto-East Cushitic	*d', *d'₁, *k'	Sasse (1979, p. 5)
Proto-East Cushitic	*b', *d', *t', *c', *j', *k', *g'	Ehret (1991, p. 273)
Proto-East Cushitic	*d', *D	Bender (2020)
Proto-Cushitic	-	Bender (2020)
Proto-Cushitic	*p', *t', *ts', *c', *k', *k^w'	Ehret (1987, p. 163)
Proto-Omoti	*p', *b', *t', *ts', *c', *k'	Bender (1988, p. 139)
Proto-Omoti	*p', *d', *ts', *s', *c', *k'	Ehret (1995, pp. 10–12)
Proto-South Omoti	*p'/*b', *t'/*d', *s'/*ts', *c', *k'/*g'	Fleming (1988, pp. 166–172)

Hudson (1989, p. 9) provides phonetic descriptions of the non-pulmonic consonants attested in HEC languages: ejectives (**p', t', c', k'**); glottalized and somewhat ingressive **d'** in Burji, with **b'** found only in Burji; glottalized and retroflex **d'** in Gedeo and Sidaama; and dialectal variants of glottalized sonorants (**'l ~ ʔl, 'r ~ ʔr, 'n ~ ʔn**) in Burji, and possibly **'l ~ ʔl** in Gedeo. These phonetic distinctions become relevant when evaluating whether **d'** reflexes represent inherited phonemic contrasts or contact-induced phonetic variation (see discussion under PHEC ***t'**).

Bender (2020, pp. 155, 138, 176) has not reconstructed any non-pulmonic consonants for Proto-Cushitic and for Proto-East Cushitic only an implosive ***d'** and, provisionally, a retroflex ***D**. Black (1974, p. 280) similarly reconstructed only a retroflex ***D** and a uvular ***q** for Proto-Lowland East Cushitic.

Early studies have used the distribution of non-pulmonic consonants to classify typological features within Cushitic languages. According to Fleming, as cited in Hudson (1976, p. 246), glottalized ejectives distinguish HEC from Lowland East Cushitic (LEC), the latter being characterized by pharyngeal consonants. This feature groups HEC with Oromo-Konsoid-Werizoid rather than other Lowland languages. Fleming (1976, pp. 306–307) attributes this HEC vs. LEC distinction to Omoti influence on HEC. The distinction between HEC and LEC is largely attributed to the influence of Omoti languages on HEC.

Fleming's claim has been the subject of an ongoing debate about whether the non-pulmonic consonants in Highland East Cushitic languages are inherited

or borrowed.¹⁹ Arvanites (1991) argues for an inherited status of **t'** and **c'** in addition to **d'** and **k'** from Proto-East Cushitic. She rejects Fleming's hypothesis of large-scale borrowing from Omotic into HEC. Sasse (1979, cited in Arvanites, 1991, p. 122) supports the inheritance hypothesis over borrowing when good correspondences can be found in EC.

Crass (2002) also argues for the hypothesis of contact. He suggests that the non-pulmonic consonants disappeared over time in many languages and were re-established later due to language contact. Crass attributes the re-establishment of the non-pulmonics in various Cushitic languages to contact with Amharic, Tigrinya, or Oromo (Crass, 2002, pp. 1686, 1688). Although claims in Crass (2002) lack evidence, they are worthy of further investigation.

The above overview reveals the historical complexity of non-pulmonic consonants in (Highland) East Cushitic languages. Scholarly approaches to their reconstruction vary considerably. That is because, unlike most of the pulmonic consonants examined in §2.1.1, the non-pulmonic consonants show substantial irregular reflexes, making the contact-inheritance analysis particularly critical for this consonant class. The present analysis applies the contact/inheritance continuum articulated in §1.5.4, which recognizes inherited proto-phonemes and contact-influenced developments as endpoints on a continuum rather than mutually exclusive categories, to be evaluated case-by-case based on distributional evidence.

Following this approach, the analysis examines each proto-phoneme through multiple lenses simultaneously: (1) regular correspondences across HEC languages backed up with wider EC context to establish inherited phonemes, (2) distributional irregularities to flag potential contact influence such as geographic clustering or irregular reflexes, and (3) areal context from external comparison to Oromo, Omotic, and other neighboring families (see Hudson, 1989, p. 8). Where correspondences are regular and evidence substantial, the pattern is treated as inherited per the parsimony principle presented in (§1.5.4), while irregular reflexes and suspicious distributions reveal contact-influenced sound changes.

Applying the probability scale (§1.5.1) and contact/inheritance continuum (§1.5.4), the evidence supports reconstructing two PHEC proto-phonemes in

¹⁹ According to Fleming (1976, p. 307) “[p]honologically... adjacent Semitic, Burji-Sidamo [HEC], and Omotic languages share highly similar sets of features, especially four or five glottalic ejectives”.

the non-pulmonic alveolar-post-alveolar series: ***t'** and ***c'** (reasonable). Both show regular correspondences across HEC languages, substantial internal supporting evidence, and promising East Cushitic parallels. The analysis examines **d'**, attested in Burji, Gedeo, and Sidaama, and analyzes it as a contact-influenced development within the **t'** correspondence set.

PHEC **t'*

The reconstruction of PHEC ***t'** is reasonable and supported by regular reflexes across all five HEC languages, initial and non-initial distribution patterns, and clear East Cushitic cognates establishing ***t'** as a potential PEC proto-phoneme. However, ***t'** exhibits remarkable reflex diversity in Burji, Gedeo, and Sidaama, where irregular **d'** and **c'** reflexes arise from intensive contact with neighboring Oromo, Omotic, and Amharic languages. These irregular correspondences represent contact-induced developments in the inherited lexicon, illustrating the contact/inheritance continuum framework established in §1.5.4.

The following analysis presents evidence sequentially: regular correspondences across different positions in a word establishing the proto-form, comparative East Cushitic evidence demonstrating inheritance, and a detailed examination of irregular reflexes with explicit contact explanations. This integrated presentation reflects the reality that the same lexical sources can exhibit both regular (inherited) and irregular (contact-influenced) reflexes across the language family.

Arvanites (1991, p. 136) extracts the non-pulmonic PEC consonant ***t'** (and ***c'**) from the reflexes of Sasse’s problematic PEC ***d’₁**. According to the available data (§1.3), **t'** occurs in approximately 10 instances in Burji, while **c'** appears in over 25. Excluding Burji, there is additional evidence of **t'** across different word positions in NHEC. Below is a summary of the attested series of correspondences, their positions, and frequency covered in the analysis:

	Position	Burji	Had	Kam	Sid	Ged	Frequency
1.	Initial, non-initial	t'	t'	t'	t'	t'	Common
2.	Initial	d'	t'	t'	d'/t'	d'	Uncommon
3.	Initial	d'	ø	ø	ø	ø	Uncommon
4.	Non-initial	ʔ/d'	ø/ʔ	ø/ʔ	ø/ʔ/d'	ø/ʔ/d'	Common (suffix)
5.	Initial	c'	t'	t'	t'	t'	Uncommon
6.	Initial	c'	t'	t'	c'	c'	Uncommon

Except for the first series, all include irregular reflexes. The slash between variants, e.g., Sidaama's initial **d'/t'**, indicates that both are attested within that series in different lexemes or even the same lexeme in the case of Sidaama. This analysis will follow the above order: starting with series 1 (with regular **t'**), then series 2–4 (with irregular **d'**), and finally series 5–6 (with irregular **c'**).

Before turning to the mixed **c'** series, I state the diagnostic used here. I reconstruct inherited PHEC ***c'** only where **c'** is the regular reflex across all five HEC languages; those sets are treated separately below (see PHEC ***c'**). Where Hadiyya and Kambaata retain **t'**, and **c'** is confined to Burji or the southern languages (Gedeo and Sidaama), I treat **c'** as an irregular development within the ***t'** correspondence set. I invoke contact when mixed HEC correspondence, a restricted southern distribution, the lack of EC parallels, and a plausible Oromo or Omotic model coincide.

Regular reflexes of PHEC **t'*

The following examples provide evidence for the reconstruction of ***t'** in initial position with regular correspondences across HEC languages.

	[351]		[308]		[269]	
PHEC	*t'uut'-	'to suck'	*t'uur-	'to milk'	*t'aʔ-,	'to extinguish'
					*t'eʔ-	'to diminish'
Burji	t'uunt'-	'absorb'	t'uur-	'to milk'	t'ay-,	'extinguish'
					t'ey-	'dry out'
PNHEC	*t'uut'-	'to suck'	-	-	*t'aʔ-,	'to extinguish'
					*t'eʔ-	'to diminish'
Had	t'uut'-	'to suck'	-	-	t'aʔa	'be depleted/ diminish'
Kam	t'uut'-	'to suck'	t'uur-	'to milk'	t'aʔ-,	'extinguish'
					t'eʔ-	'be depleted'
Sid	t'uut'-	'to suck'	t'uur-	'to milk'	t'eʔ-	'be depleted'
Ged	t'uut'-	'to suck'	-	-	t'eʔy-	'be incomplete'

Set **[351]** presents ***t'uut'-** 'to suck' reconstructed from regular correspondences in HEC languages. Certain sound changes are notable within these sets. For instance, in **[351]**, the extra **-n-** in Burji **t'uunt'-** is an inserted sound (§3.4.3). The common vowel variation before the glottal stop in **[269]** is discussed in §2.2.2.1.

In non-initial position, evidence for PHEC ***t'** from all HEC languages comes from the following reconstructions:

PHEC [304] **wit'*- 'scatter' → Burji *wat'ar-*, *wot'ar-*; PNHEC **wit'*- 'to sow' → Had *wit'-*, Kam *wit'-*, Sid *wit'-*, Ged (*wiisa*).

PHEC [240] **k'ort'umm-* 'fish' → Burji *k'urt'ummee*; PNHEC **k'ort'umm-* → Had *k'urt'ume?*, Kam *k'urt'ummeema-ta* / *k'urt'ummee-ta*, Sid *k'ult'u?me*, Ged *k'ult'u?me*.

PHEC [10] **hat't'if y* 'to sneeze' → Burji *hat't'iffee-d'*- (cf. *hattifeed'*-, *hattifeys-* (Roba Dame & Wedekind, 2008, p. 37)), *hat'if-*; PNHEC **hat't'if y* → Had *hat'iffuy-*, *hat't'ef-*, Kam *hant'iff-a?*, Sid *hant'iff-i?r-*, *hant'iffo*.

In set [304], the Gedeo form *wiisa* may contain a causative **s**, with double **ii** possibly indicating the reduction of **t's** clusters. The onomatopoetic nature of set [10] might have contributed to irregularity in the quantity and quality of some of the correspondences.

Additional initial and non-initial evidence of **t'** appears within NHEC in examples such as PNHEC [55] **t'ut'*- 'black ant', [744] **fit't'*- 'to disperse', [52] **hit'*- 'grass (sp.)', [613] **t'een-* 'rain', [661] **t'umm-* 'peace', [730] **t'a?m-* 'to ask, visit', [224] **t'ur-* 'dirt of body/clothing', and [75] **t'e?m-* 'to taste good'.

The regularity of correspondences across initial and non-initial examples supports the reconstruction of PHEC ***t'**, which is further supported by some EC parallels with a corresponding **t'**, albeit with some irregularities in the corresponding sounds other than **t'**:

PHEC [222] **t'ok-* 'to burst': Gaw *t'uh-* 'to sprout', Dah *t'oʕ-* 'to pain/cut'.

PHEC [10] **hat't'if y* 'to sneeze': Arb *het't'iris-*, Gaw *ħat't'is-*, Bay *hant'irfe* 'sneeze'.

PNHEC [55] **t'ut'*- 'black ant': Yaaku *t'oro?te* 'driver or safari ant', Dah *t'at'e* 'small ant', Oro *t'iint-* 'maggot'.

PHEC [250] **t'iib-* 'to push': Oro *t'ibiira* 'twist (like in making ropes)', Gaw *t'iip-* 'to squeeze', Dah *t'uʊʕ* 'to press/throttle', *t'uub* 'squeeze'.

Beyond HEC, according to Ehret (1987, p. 34), Yaaku **t'** and Oromo **d'** are reflexes of PC ***t'**, while Oromo words with **t'** are possibly borrowings.²⁰

Certain HEC examples exhibit one parallel in a neighboring language, such as PHEC [351] **t'uut'*- 'to suck', with a comparable stem in Oromo *t'uut'a* 'suck', and HEC [269] **t'aʔ*- 'to extinguish', maybe with Gawwada *dəʔ*- 'to dry up and crack'. PHEC [304] **wit'*- 'to sow' might have been transferred into Bayso as *wit'*- 'sow, seed' and it is comparable to Hamar *yiiit'* 'sow, planting seeds by scattering them on the earth' (Petrillino, 2016, p. 318) and Sheko *yiiits'* 'to sprinkle' (Hellenthal, 2010, p. 975). These HEC-only parallels point to external (non-HEC) origins or HEC innovations transferred into the neighboring languages.

Six of the reconstructed roots with PHEC ***t'** have no comparable parallels in EC or the neighboring Omotic/Amharic languages: PHEC [308] **t'uur*- 'to milk',²¹ [744] **firt'*- 'to disperse',²² and PNHEC [52] **hit'*- 'grass (sp.)', PNHEC [613] **t'een*- 'rain', [661] **t'umm*- 'peace', and [730] **t'aʔm*- 'to ask, visit'. In these cases, the absence of parallels suggests that some HEC roots may be innovations or borrowings from an unknown source.

A few of the above HEC roots have parallel examples with a corresponding **t'** in the neighboring languages, particularly in Oromo:

HEC				Parallels			
Gedeo	[10]	het't'if-at-	'sneeze'	Oromo	hat't'iffatu	'sneeze'	
PHEC	[240]	*k'ort'umm-	'fish'	Oromo	k'urt'ummii	'fish'	
PHEC	[351]	*t'uut'-	'to suck'	Oromo	t'uut'a	'suck'	
				Amharic	t'ut	'breast'	
PHEC	[269]	*t'aʔ-	'extinguish'	Wolaytta	t'ay-	'to get lost'	

In [10], Gedeo *het't'if-at-* exhibits a typical Oromo sound change, ***s** → **f**, which might be a result of later contact influence. PHEC [240] **k'ort'umm*- 'fish' is another Oromo loan from *k'urt'ummii* 'fish' (Sasse, 1982, p. 129; see Hudson,

²⁰ Sasse (1979, pp. 29–30) proposes a medial PHEC ***t'** but for different reflexes, namely Hadiyya **ʔ**, Sidaama and Gedeo **d'**. These reflexes rather include various suffixes discussed in Ch. 4.

²¹ PHEC [308] **t'uur*- 'to milk' might be comparable with Burunge *turuʔud* and Alagwa *turuʔut* 'sip, suck' in the West-Rift Cushitic languages (Kießling & Mous, 2003, p. 275).

²² PHEC [744] **firt'*- 'to disperse', [721] **fir*- 'relative', and [162] **fic'*- 'to scatter' might be related. Kießling & Mous (2003, p. 108) reconstruct PWR **fiits'*- 'sweep, level', which they link to Sidaama **fit* 'comb' and Somali *fid* 'comb'. Those are rather linked to PNHEC **fil-ʔ*- 'comb' in this study. Lamberti & Sottile (1997, p. 357) interpret Zaysse *pits'*- 'sweep' to include the benefactive suffix ***-ʔ**- that has caused the glottalization of the dental in "the old West Cushitic stem" **bit-*, after which ***t'**- became **-ts'**-.

1989, p. 8). However, it might be a very recent or a very old loan because it is found all over HEC. PHEC [351] **t'uut'*- 'to suck' may be an original PHEC root, but it is also found in Oromo *t'uut'a* 'suck' and Amharic *t'ut* 'breast' (see Burji *t'uunt'*- 'to suck' and Zaysse *ts'unts'*- 'suck out' (Sasse, 1982, p. 183)). The same applies to PHEC [269] **t'aʔ*- 'to extinguish' with its parallels in Wolaytta *t'ay*- 'to get lost' and other Omoto languages, raising questions about the direction of transfer. Since the latter two are regularly reconstructed for PHEC, they must be inherited PHEC roots.

Irregular d' reflexes of PHEC *t'

The preceding analysis establishes the reconstruction of PHEC **t'* based on regular correspondences in different word positions and East Cushitic cognates. However, Burji, Gedeo, and Sidaama exhibit irregular **d'** reflexes in specific lexical sets, requiring detailed examination.²³ Such irregular patterns correlate with documented Oromo contact and do not represent a separate proto-phoneme.

An example of the irregularity is that Burji initial **d'** may correspond to either **d'** or **t'** in Sidaama, as illustrated in the following examples:

²³ The synchronic description of **d'** ranges between implosive and sometimes a retroflex in Gedeo, Sidaama, and Burji (Hudson, 1976, p. 248; Bender, 2020, pp. 116–117; Treis, 2008, p. 27). Yri & Pepper (2019) state, in Sidaama, the "...implosive can be described as having an oral and a glottal gesture [glottalized retroflex stop (Hudson, 1976, p. 248)]. Following a liquid or a nasal, it undergoes (underwent) reduction, i.e., loss of its oral gesture and metathesis, probably in that order. The result was a cluster of glottal + liquid/nasal". This reduction hypothesis is rejected as explained in §§4.2.3, 3.3.2.

	[214]		[250]		[128]	
PHEC	*t'aban ʔ-s	'to slap'	*t'iib	'to push'	*t'eeb ʔ	'be thirsty'
Burji	d'abanaas	'to slap'	d'iib-	'to push'	d'aybood'-	'be thirsty'
PNHE	*t'aban ʔ-	'to slap'	*t'iib-	'to push'	*t'eeb ʔ	'be thirsty'
C	s-					
Had	t'aban-s-	'to slap'	t'iib-	'to push'	t'eebeʔ-	'be thirsty'
Kam	t'aban-s-	'to slap'	t'iib-	'press'	t'eebeʔ-	'be very thirsty'
Sid	d'aʔa assa	'to slap'	t'iibi	'to push'	-	-
Ged	d'abanaas	'to slap'	d'iib-	'push'	d'eebota	'thirst'
cf. Oro	d'aana	'beat on'	d'iib-	'push'	d'eebotee	'thirsty'

Examples [214], [250], and [128] are the only examples supporting the putative initial PHEC ***d'**, though the Oromo parallels reveal the source of the irregular **d'** reflexes in Burji, Gedeo, and Sidaama. Set [214] provides the strongest evidence in favor of a PHEC ***d'** because it has the reflex **d'** in three HEC languages. The other two examples [250] and [128] lack a supporting **d'** reflex in Sidaama. Hence, they are all reconstructed to P(N)HEC ***t'**, analyzing the **d'** reflex as irregular, attributed to Oromo influence on the three languages and supported by further evidence below.

Overall, the sound **d'** is predominantly found in the southernmost language, Burji. It appears less frequently in Gedeo and is least common in Sidaama. In Sidaama, initial **d'** is sometimes in free variation with **t'** as in the examples below.²⁴

	t'		d'	Gloss
[417]	t'ibba	~	d'ibba	'disease'
[549]	t'aam-	~	d'aam-	'to give a message to'
[201]	t'agge	~	d'agge	'story, miracle'
[734]	t'ara	~	d'ara	'false'
[159]	t'aawa	~	d'aawa	'to erect, establish'

Wedekind (1980, p. 149) identifies the potential sources of **t' ~ d'** variation by noting that [201] *t'agge* is from Northern Sidaama, while *d'aggi* 'news' is from Southern Sidaama varieties.²⁵ Thus, the Southern Sidaama forms with a

²⁴ Hudson (1976, p. 248) describes **d'** as being primarily glottalized in both Gedeo and Burji, somewhat less so in Sidaama, and entirely absent from Hadiyya and Kambaata. This description of **d'** aligns with its distribution from south to north.

²⁵ Note that the above [201] examples are from a different source, hence the slight variation in transcription.

corresponding **d'** are the most influenced by Oromo/Burji (see Sasse (1979, p. 30)). Further south, Oromo contact with Gedeo is evident in plenty of loanwords, e.g., [346] Gedeo *d'uufu* 'fart' (← **d'uus-*) since **s* → *f* is an Oromo innovation. Further examples with an irregular initial **d'** attested in Burji, Gedeo, and Sidaama only are found in [549] (Oromo loan), [734], and [78].

In a rare secondary series, Burji has **d'**, while other HEC languages display the reflexes **ø/?** in initial position. This series appears in only two or three comparanda, suggesting a potential fossilized reflexive prefix (§4.2.3).

Sasse (1982, p. 62) reiterates Hudson's (1976, p. 272) connection of the verb *d'ayb-ad'* 'be warm' to the common EC stem *iib-* 'be warm' in Burji and interprets the initial **d'** as a fossilized middle voice prefix, similar to that in *d'irk'* 'give back'. Hayward (1984, p. 84) confirms that "the only languages having prefixed root extensions are Saho, °Afar and Burji... when due consideration has been given to morphophonological alternations the °Afar and Saho forms can only be accounted for by means of an underlying prefix with **t**, while the Burji forms... clearly contain a prefix with a **df**....". This phenomenon has been observed in examples like:

PHEC [298] **a/?* 'to be useful' → Burji *d'al/?* 'to be useful'; PNHEC **e/?* → Had *e/?* 'be better, improve', Sid *e/-* 'to grant, make donation, use, be useful to', Ged *e/-* 'be better, improve'.

PHEC [306] **y ebb-* 'be(come) hot/warm' → Burji *d'aybad'*; PNHEC **y ebb-* → Had, Kam, Sid *iibb-*, Ged *eebb-*.

The above evidence, in addition to Burji *d'irk'*, (intr) and *rik'k'* 'return' (tr), supports the observation that there are reflexives once formed with a prefix in Burji (see Sasse, 1982, p. 65). Further reiteration of Hudson's observation is found in Hetzron (1980, p. 95), who states that "[n]ormally in Cushitic derivational affixes are postverbal. Yet in Saho and Afar, prefix-conjugated verbs have prefixed derivational affixes... G. Hudson (1976:272) has shown that even Burji ... has some vestiges of original derivational prefixes, now lexicalized, e.g. *ḍirk'* 'return (intr)', vs. *rikk'* 'return (tr.)' with a reflexive prefix (which is now a suffix *-(a)ḍ*).". Hetzron provides further evidence of fossilized prefixes "[i]n Rendille, the verb *-tVḥ* 'become' may have been formed as a reflexive (medium) of *-daḥ* 'say' ... and the root *-sil(l)* 'lay down' could be a causative of *-al(l)* 'be placed, exist, remain'. The reconstructed PHEC reflexive is **-?* (§4.2.3.1), which means Burji **d'** in the prefix is still unique and irregular within HEC, suggesting Oromo as the source (§4.2.3.4). An

alternative potential source for the NHEC reflexes could also be the derivational verbalizer/support verb *ʔ since [306] *y *ebb-* ‘be(come) hot/warm’ above also includes the other verbalizer/support verb *y (§§4.2.4, 4.2.5).

The above evidence signifies the distinction between Burji and NHEC based on both C₁ and V₁. Additionally, it raises questions such as how many other cases there are of such prefixes beyond Burji²⁶ and how much that contributes to the idiosyncratic correspondences.

Looking at the reflexes in non-initial position, HEC languages demonstrate an even more complex set of reflexes with irregular **d'**, which reconstruct to PNHEC and PHEC *ʔ. Such analysis is contrary to Bender (2020, pp. 116–117) and Hudson (1989, p. 7), who reconstruct a PHEC ***d'** based on its presence in Burji, Gedeo, and Sidaama, with Hadiyya and Kambaata showing ʔ as a reflex. Those reflexes are found mainly in a series of non-initial correspondences with irregular reflexes, as in the following examples:

	[348]		[168]	
PHEC	*c'iiʔ-	‘bird’	*heeʔ-	‘to exist, to live’
Burji	c'iid'aa	‘bird’	(heʔ-, yed'-)	‘here it is!, exist’
PNHEC	*c'iiʔ-i-ff-	‘bird’	*heeʔ-	‘to exist, to live’
Had	c'iiʔ-cco	‘bird’	heeʔ-	‘live, exist’
Kam	c'iiccu-	‘bird’	heʔ-	‘live’
Sid	c'eʔicco	‘bird’ ²⁷	heed'-	‘exist, be’
Ged	c'iʔa	‘bird’	hed'-	‘exist, be there’

In comparanda [348], **d'** occurs in Burji only, whereas in [168] it is found in Sidaama and Gedeo examples (the Burji cognate is doubtful). Sasse (1979, pp. 29–30) proposes a medial (non-initial) PHEC ***t'** for the correspondence set of Hadiyya ʔ, Sidaama, and Gedeo **d'**, which resembles the correspondences in [168]. However, the glottal stop in [168] involves retained and attached reflexive/derivational glottal stops, while the **d'** is a reflexive/derivational suffix in Sidaama and Gedeo (§§4.2.3). In [348], Burji *c'iid'aa* ‘bird’ is best treated as an isolated Burji irregular non-initial reflex of *ʔ rather than as evidence for reconstructing medial PHEC ***t'** or ***d'**.

²⁶ A comparable instance is found in Oromo [43] *d'uga* ‘drink’ vs. PS **uk-* ‘to drink’.

²⁷ Sidaama *c'eʔicco* ‘bird’ is contractable into *c'icco* (Gasparini, 1983, p. 56).

60 A Lexical and Phonological Reconstruction of Highland East Cushitic

There are many more examples of non-initial **d'** found on derived verbs; thus, they are independently added suffixes (§§4.2.4, 4.2.5) as the following correspondences demonstrate:

[256]	Ø	-ʔ	-t	-d'	Gloss
PHEC	-	*ham ʔ			'to slander'
Burji				ham-ad'-	'slander'
Had	ham-				'slander'
Kam		heʔm-			'slander, gossip'
Sid	heme	heʔm-			'backbiting', 'slander'
Ged			ham-at-		'slander'
cf. Oro			hammaataa	hamad'd'a	'bad off', 'slander'
[450]	Ø	-ʔ	-t	-d'	Gloss
PHEC	*jaal-				'be shy, be ashamed'
Burji		ja(a)lʔ-		saalfad'-	'be shy', 'be embarrassed'
Kam	saal-				'be shy, ashamed'
Sid	saal-		saalfat-		'be shy, ashamed'
Ged			saalfat-		'be shy, ashamed'
cf. Oro				saalf-ad'd'a	'be shy, ashamed'
[1030]	Ø	-t	-d'		Gloss
PHEC	*huk'-				'be(come) thin'
Burji				huk'k'ad'-	'to get thin'
Ged		huk'k'ata			'be slim'
cf. Oro		huk'k'at-	huk'k'ad'd'a		'thin', 'become thin'

The above examples show that the correspondence sets include at least **Ø**, **ʔ**, **t**, and **d'**. Again, Oromo influence is noticeable where Burji has **d'** and Gedeo and Sidaama have a corresponding **t** both of which are suffixes given the formations provided in the Oromo dictionaries (Gragg, 1982; Tilahun Gamta, 2004; Stroemer, 2001).²⁸ The other HEC languages have corresponding

²⁸ Oromo **-t-** could be interpreted as a variant of the middle voice verb extensions of Borana (Marsabit) and Waata dialects reported by Stroemer (1987, pp. 153, 155) that produce variants **-ad'd'-** (for first-person singular) and **-at-** (elsewhere) with occasional variation in vowel quality. It might also be a different derivational suffix like the one found in HEC (§4.2.5.3).

underived stems with $\emptyset$, such as in Hadiyya, Kambaata, and Sidaama, and stems with the reflexive/derivational $-\text{ʔ}$ in Hadiyya and Kambaata ([450]).

While Burji $-\text{d}'$ is a loan from Oromo, Sidaama, and Gedeo $-\text{t}$ is likely a contact-induced retention of an agentive denominative (cf. §4.2.5.3). They were likely already cognates in Gedeo and Sidaama and have been influenced by the Oromo words ending in $-\text{t}$.

See also [497], [177], [340], and [698] for further examples.

The irregular correspondences in initial and non-initial positions, the phonological processes at morpheme boundaries (§3.3.1.5), and the influence of Oromo on certain Sidaama and Gedeo forms and loans in Burji challenge any proposal of a consistent reconstruction for $*\text{d}'$ for PHEC.

The wider East Cushitic context reveals greater complexity but could be synthesized as the most common correspondences of PHEC $*\text{t}'$ are probably retained t' in Southern East Cushitic (Yaaku-Dullay), and innovated d' in the Lowland East Cushitic branch. Since the EC parallels with a corresponding t' have been discussed above, the following paragraphs will be dedicated to the ones in which the regular and irregular reflexes of PHEC $*\text{t}'$ correspond to d' (realized as d' , d , or q) in the rest of EC.

PHEC [222] $*\text{t}'\text{ok-}$ 'to burst': Oro *d'oʔa* 'burst, explode', Dir *dēh* 'grow', Mos *he-dēhe* 'sprout', Kon *dēh-* 'sprout', Afar *dokʕ-* 'bruise, dent', Arb *d'oʔ-* 'burst'.

PNHEC [55] $*\text{t}'\text{ut-}$ and PHEC $*\text{t}'\text{uut-}$ 'black ant': Dir *dūuddéṭ* 'small black ants, black beads', Mos *dūde-ta* 'ant', Saho *quuḍe* 'ant', Afar *dūuné* 'ants', Ren *qūuḍe* 'black ant(s)', Som *quurcaso* 'ant species reddish orange in color', *quuqi* 'a kind of louse', Maay *quuri* 'ants', Boni Kilii *d'óóí* 'ant sp. (small)'.

PNHEC [75] $*\text{t}'\text{eʔm-}$ 'to taste good': Oro *d'amd'ama*, 'taste lightly', *d'an-d'am-* 'to taste', Saho *qafame* 'to taste', Afar *dafame* 'to taste', Bay *t'am-*, Som *qadami* 'taste', Maay *deqan* 'taste', Boni *d'id'imiy-* 'taste', Das *d'ád'am* 'taste'.

PHEC [214] $*\text{t}'\text{aban } \text{ʔ}$, $*\text{t}'\text{aban } \text{ʔ-s-}$ 'to slap' ← PHEC [106] $*\text{t}'\text{aban-}$ 'side, cheek': Kon *dapn-a* 'side (of face, body, mountain, river)', Oro *d'amn-a* 'jaw', Saho *qaban* 'ceek, side', Afar *dában* 'cheek, side', Ren *qában*

62 A Lexical and Phonological Reconstruction of Highland East Cushitic

‘side’, Som *qaban* ‘cheek’, Maay *qabaŋk* ‘cheek’, Boni *d’áŋ* ‘side, direction’.

PHEC [250] **t’iib-* ‘to push’: Kon *dīip-p-* ‘grasp, squeeze’, Som *qīiqīib-* ‘bring thighs together to close crutch [crotch]’, Saho *qībro* ‘catching, hold’, Afar *dībbaanise* ‘connect together’, PEC **d’iib-* ‘squeeze, press’ (Sasse, 1979, p. 29).

In [222], the reflexes in Sidaama and Gedeo show irregular **d’** in comparison to the rest of HEC.

Some of the above examples have variants in both HEC and wider EC recorded with a consonant other than **d’** such as **c’**, **j’**, **ɟ**, **c**, **t**, and **d**.

[222] Burji *tokkaʔ-* ‘to leak/drop’, *dar-ʔ-* ‘burst’, Sid *taʔi* ‘burst, crack, explode’, Ged *doʔ-* ‘burst’.

[55] Burji *c’uuc’c’ee* ‘black ant’ and Had *c’uuma* ‘big black ants’, Boni Baddey *j’úúrə*, Boni Bireeri *c’úúrə* ‘ant sp. (small)’, Bayso *tunce* ‘red ant’.

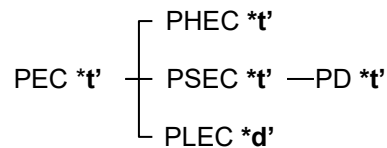
[75] Sid *c’oommi* ‘taste good’, Boni *iʔimiy*, Afar *taam-* ‘to taste’, Ren *camcama* ‘(to) taste’, Das *ɟam* (adj) ‘taste, strong taste’, Dah *tem-* ‘to taste’.

[250] Dir *c’iip-* ‘squeeze’, Mos *c’iib-eeca* ‘squeeze, wring out’, Saho *dufuwe* ‘to push’, Afar *duuf-* ‘forceful push’.

Because these forms coexist with **d’** reflexes in the same language, I do not treat them as evidence for a single regular East Cushitic correspondence. At most, they indicate mixed dialectal variants or secondary lexical interference. In [55] PHEC **t’uut’-* ‘black ant’, Burji and Hadiyya **c’** diverge from reconstructed ***t’**; however, Boni parallels also have **c’**. The variation in the reflexes of this set might be due to references to different species of ants, as the glossing suggests, or multiple contact sources. The second most common irregular EC correspondence of PHEC ***t’** is **c’**, such as in [55], and will be addressed in detail in the next series of irregular correspondences.

In Lowland EC, the most common sound is **d’**, which corresponds to **t’** in HEC, SEC, and Dahalo languages. Hence, PHEC ***t’**, PSEC ***t’**, and PLEC ***d’**

reflexes support the reconstruction of PEC *t' (see PC *t' in Ehret (1987, p. 34)) according to the traditional subgrouping:



In conclusion, **d'** is a phoneme in Burji, Gedeo, and Sidaama; however, it is not reconstructable to P(N)HEC ***d'** due to the irregular correspondences and evident influence of Oromo over these three languages. Thus, initial **d'** in Burji, Gedeo, and Sidaama (i.e., from series 2–3 (see *Irregular d' reflexes of PHEC *t'*)) results from independent contact with Oromo, while the correspondences in non-initial position from series 4 (i.e., the reflexive/derivational suffix) developed from contact-induced change and via a regular fusion process of **r** and **ʔ** in Sidaama and Gedeo, as discussed in the next chapter in §3.3.1.5.

Irregular c' reflexes of PHEC *t'

The rest of the irregular reflexes of PHEC ***t'** include a corresponding irregular **c'** as already seen in some of the EC parallels above. Therefore, the remaining paragraphs will elaborate on those correspondences where Burji has **c'**, Gedeo and Sidaama have **c'** or **t'**, and Hadiyya and Kambaata have **t'** (series 5–6).

The comparative sets [164] and [202] below provide the only consistent correspondences where both Gedeo and Sidaama exhibit initial **c'** in addition to Burji:

64 A Lexical and Phonological Reconstruction of Highland East Cushitic

	[164]		[202]		[743]	
PHEC	*t'uf-	'to close'	*t'all-	'only, alone'	-	-
Burji	c'ukk-	'to close'	c'alla	'only, alone'	-	-
PNHEC	*t'uf-	'to close'	*t'all-	'only, alone'	*t'ork'	'to chew'
Had	t'uf-	'to close'	t'alla	'only, alone'	t'orok'-	'be drilled'
Kam	t'uf-	'to close'	t'alla	'only, alone'	t'ork'o	'hole'
Sid	c'uf-	'to close'	c'alla	'only, alone'	c'ork'a	'to chew'
Ged	c'uf-	'to close'	c'alla	'only, alone'	c'ork'a	'to chew'
Cf. Oro	c'ufa	'close'	c'alla	'alone'	-	-

Other than [\[164\]](#), [\[202\]](#), and [\[743\]](#), the relevant sets show **t'** in Hadiyya and Kambaata, and usually also in Sidaama and Gedeo; only Burji consistently has **c'**. I therefore treat the Burji **c'** as a language-specific secondary reflex, not as evidence for reconstructing PHEC ***c'** for the set as a whole, since the reflexes of PHEC ***c'** are **c'** throughout HEC as discussed in the next section. This is evident in the following initial correspondence. This is evident in the following initial correspondences:

	[224]		[613]		[55]	
PHEC	*t'ur-	'dirt'	*t'een-	'rain'	*t'ut'-	'black ant'
Burji	c'ure	'rubbish'	c'eena	'rain'	c'uuc'c'ee	'black ant'
PNHEC	*t'ur-	'dirt'	*t'een-	'rain'	*t'ut'-	'black ant'
Had	t'ura	'dirt'	t'eena	'rain'	-	-
Kam	t'uri	'dirt'	t'eena	'rain'	t'ut'i-	'little ant'
Sid	t'ure	'dirt'	t'eena	'rain'	t'ut'e,	'big black ants'
Ged	t'ure	'dirt'	t'eena	'rain'	c'uunc'a	'small black ants'
Cf. Oro	t'urii	'dirt'	-	-	t'iint'-	'maggot'

The correspondence series with irregular **c'** is found in initial position only (cf. *Regular reflexes of PHEC *t'*).

Except [\[202\]](#)²⁹ and [\[613\]](#), the other four examples above have parallels in the rest of the EC languages, with greater diversity in the reflexes, including the ones attested above and others, as presented below:

²⁹ PEOT has **kald'ay* 'alone, only' but the correspondences are also problematic when compared with PHEC [\[202\]](#) **t'alla* 'only, alone'.

[164]	t'	d'	c', ɸ, j	Gloss
PHEC	*t'uf-			'to close'
Oro			c'ufa	'close'
Dir		ɗuuɸ-		'cover, put a lid on'
Mos		ɗuɸ-		'cover'
Kon		ɗuf-		'close, enclose'
Bay			c'uf-	'close'
Das			ɸobónoɸ-	'cause it to close'
Som			jufee	'insert a plug'

Two languages show a corresponding **t** and **d** for [164] Elmolo *tibʔe* 'to bury' and Yaaku *deps* 'to close, to shut, to cover, to bury, to plug'. Sasse (1979, p. 20, 31) reconstructs PEC **d',uf-* 'close, shut' for the above correspondence.

[743]	t'	d'	c', c, j	Gloss
PNHEC	*t'ork'			'to chew'
Burji		d'okkomm-		'to chew (roasted grain or cane)'
Ts'am			č'óx-	'milk'
Gaw			c'oh-a	'to milk'
Yaaku			c'aqau	'to chew'
Elm			cak-	'to chew'
Som			jaq	'suck something'

The above examples are problematic, including Burji [743] *d'okkomm-* 'to chew' and the EC parallels, which lack the PHEC **r*.

[224]	t'	d'	ɸ, ɹ	Gloss
PHEC	*t'ur-			'dirt of body, clothing'
Oro	t'urii			'dirt, filth, dirt of body, clothing'
Dir	t'ur-á			'dust'
Arb	t'uré			'dirty thing'
Afar		ɗorók		'dirt'
Kon			ɸur-eeta, ɹur-	'dirt, butter'

Different correspondences are found in Mositacha *ts'ura* 'dust' and Yaaku *durug* 'dirt'. Again, Sasse reconstructs PEC **d',ur-* 'dirt' with a problematic **d',* for such instances of **d'** corresponding to **c'**, **ɸ**, or **t'** (Sasse, 1982, p. 49; Sasse 1979, p. 29, 31). Such irregularity rather suggests this root may have undergone multiple phonetic shifts due to contact.

³⁰ According to Kießling & Mous (2003, p. 286), PWR **t'up* 'smash' is the cognate of PEC **d',uf-* 'close, shut', which is reconstructed here as PHEC [164] **t'uf-* 'close'.

[55]	t'	d', d	c', j'	Gloss
PHEC	*t'ut'-			'black ant'
Oro	t'iint'-			'maggot'
Yaaku	t'oroʔtɛ			'driver or safari ant'
Dah	t'at'e			'small ant'
Dir		duuddéɪ		'small black ants'
Mos		dude-ta		'ant'
Saho		quude		'ant'
Afar		duuné		'ants'
Ren		quude		'black ant(s)'
Maay		quuri		'ants'
Som		quudɪ		'a kind of louse'
Boni		d'óói	j'úúra, c'úúra	'ant sp. (small)'

See Bayso [55] *tunce* 'red ant' and PWR *ts'ilaʔu 'small black ants (sp.)'.

Yaaku parallels in [224] and [55] still support a PHEC *t'. However, the evidence in the rest of SEC (Dullay) is not in consistent with PHEC *t', this time as far as the parallels in [743] are concerned. The same sets show a consistent Somali correspondence j.³¹ Excluding [55], d' is a less common correspondence compared to the previous sets.

Due to such irregularity, Sasse reconstructs the parallels of two of the above PHEC stems, i.e., [164] *t'uf- 'to close' and [224] *t'ur- 'dirt of body, clothing' with a problematic initial PEC *d'₁, from which Arvanites (1991, p. 136) extracts two PEC non-pulmonic consonants *t' and *c'. For the present argument, however, the decisive issue is internal HEC reconstruction rather than PEC. At the PHEC level, these sets are better assigned to *t' than to *c', because their HEC correspondences are mixed. Additionally, for the regular PHEC *c' sets discussed below, c' is retained across all five HEC languages.

The areal perspective might provide some answers for the potential sources of the irregular corresponding c' in HEC. Contact between these languages, Oromo, and Omotic languages is evident in certain cases. For instance, Burji, Gedeo, and Sidaama c' may be due to the heavy influence of Oromo, which explains why most of the cognates show a regular reflex reconstructable to PNHEC *t', where the irregular c' is isolated and mainly in Burji.

³¹ Somali j is the regular reflex of PEC *c' (Arvanites, 1991, pp. 158–159).

	Oromo	→ Burji	→ Gedeo	→ Sidaama	Gloss
[164]	c'ufa	-	c'uf-	c'uf-	'close'
[202]	c'alla	c'all-a	c'alla	c'alla	'only, alone'
[25]	c'ilee, c'ila-ttii	c'illee	c'ilaate, c'ila-tte	c'ilaatte	'coal'
[13]	hanc'ufa	-	hanc'ufa	-	'saliva'

Beyond Cushitic, Bender (1988, p. 139) and Fleming (1988, p. 171) reconstruct Proto-Omotiic **ts'* and Proto-South Omotiic **s'* or **ts'*, respectively. Thus, as an alternative to the above retention hypothesis, the sources of some Hadiyya/Kambaata *t'* and Burji *c'* might be Omotiic *ts'* and *c'*, respectively, as in the following cases:

Omotiic		→ Burji	→ Had	→ Kam	Gloss
Wolaytta <i>t'alaala</i> ³²	[202]	-	<i>t'alala</i>	<i>t'alla</i>	'only, alone'
Gofa <i>ts'alaala</i>		<i>c'all-a</i>			
Malo <i>c'uc'c'e</i>	[55]	<i>c'uuc'c'ee</i>	-		'louse', 'little ant'
Shinassha <i>ts'uts'a</i>				<i>t'ut'i</i> - ³³	'louse', 'black ant'
Wolayta <i>t'ugunta</i>	[267]	<i>t'unga</i>	-	-	'finger(nail)'
Bench <i>ts'ung</i> ³⁴					'claw'

The above examples include two of the four sets with irregular *c'*. Thus, if those two represent borrowings into HEC, the two reconstructions with PHEC **t'* (**t'all-* 'only, alone' and **t'ut'* 'black ant') would no longer be valid. However, if they are inherited HEC words that were adapted due to contact, then the original initial consonant was **t'*, as retained in Hadiyya and Kambaata.

The current dominant contact language is Amharic. Amharic loanwords with *s'/ts'* are transferred into HEC with *t'*. For instance, Treis (2008, pp. 24–25) confirms that in Kambaata "*/t'/* ... regularly replaces the Amharic ejective fricative */s'/* in loanwords", e.g.,

³² Lamberti & Sottile (1997, p. 529) propose the following changes: **d'y-* → **t'-* → *ts'-* for Gofa *ts'alaala* 'only'.

³³ In addition to Sidaama and Gedeo *t'ut'e* 'black ants'.

³⁴ Further parallels of [267] in Omotiic include Koyra *ts'ugun-tse* (Sasse, 1982, p. 183), Shinassha *ts'ungu-ttsa*, Shekko *c'ungit'o*, and Malo *s'ungu-tsa* (Lamberti & Sottile, 1997, p. 528), all meaning 'fingernail'.

Amharic *mäs'haf* 'book' → Kam *mat'aafä* (Treis, 2008, pp. 24–25), Burji *mat'aafi*, *mas'aafi* 'book' (Roba Dame & Wedekind, 2008, p. 60).

Ethiosemitic **s'-h-f* 'write' (Lamberti & Sottile, 1997, p. 475) → Amharic *s'afä*, *t'afä* 'write' → Burji *t'aaf-*, Kam *t'aaf-*, Sid *t'aaf-*, Ged *t'aaf-* (Hudson, 1989, p. 171), Oro *t'aaf-*, *c'aaf-*, and *s'aaf-* 'write' (Lamberti & Sottile, 1997, p. 525).

Amharic *s'omä* 'fast' → Burji *t'oom-*.

Amharic *mäs'haf* 'book' is related to the Ethiosemitic stem **s'-h-f* 'write' (Lamberti & Sottile, 1997, p. 475). Oromo variants *t'aaf-*, *c'aaf-*, and *s'aaf-* 'write' are dialectal variations (Lamberti & Sottile, 1997, p. 525). The existence of variants with *s'* in Burji and Oromo indicates the recency of borrowing since *s'* is not a phoneme in either.

Amharic influence is ruled out for the sets under examination because of the lack of parallels, the lack of evidence to explain Amharic's contribution of the irregular *c'*, and the attested recency of contact with HEC. Therefore, if there are irregular *c'* and *t'* from multiple contact sources – Omotic languages (Lamberti & Sottile, 1997, pp. 253, 255, 258, 259, 260, 263) and Oromo – these scenarios must be proposed set by set after careful investigation.

Since Burji exhibits a tendency toward palatalization (§3.2.1), could we propose a palatalization rule **t' → c'* to explain its irregular reflexes? There is no regular environment to justify all the occurrences of *c'* in Burji.

[224] *c'ure* 'rubbish', [164] *c'ukk-* 'close with lid, to cover', and [55] *c'uuc'c'ee* 'black ant'

[18] *c'eeji* 'blood' and [613] *c'eena* 'rain'

[270] *c'ibba*, 'hundred'

In the above examples, Burji *c'* is found before *u* in [224] and [164], *e(e)* in [18] and [613], and *i* in [270]. Even if one counts *u* amongst the sounds that trigger palatalization, there are examples with a non-palatalized *t'* before *u* in Burji [267] *t'unga* 'finger, claw', despite the environment being similar to that of [224] *c'ure* 'rubbish' and [164] *c'ukk-* 'close with lid, to cover'. Additionally, there is a palatalized *c'* without a palatalization environment, e.g., [202] *c'alla*

‘only, empty’. Thus, palatalization would not explain the irregular Burji correspondences. However, “harmony” can be proposed as an alternative hypothesis for some cases (and it is interesting that these are cases where a trigger exists for the previous palatalization hypothesis), such as in the reflexes of PHEC [18] *t’iig- ‘blood’:

Burji *c’eeji*, Gaw *c’eeq-t-e*, Ts’am *c’egde*, Boni Killi *c’iik* ‘blood’, Kam *k’egu*, *k’egi*, Gollango *k’eq-te* ‘blood’ (Bender, 2020, p. 2).

A simpler explanation of the above examples is that they borrowed from each other, but they remain irregular within the broader EC context.³⁵ Therefore, via obstruent harmony,³⁶ Kambaata and Gollango might have assimilated C₁ t’ to the velar in C₂, while Burji might have assimilated C₁ t’ to the palatalized g → j in C₂ (§3.2.1). Then, the initial k’ triggered the glottalization harmony of *g to q in Gawwada *c’eeq-t-e* and Gollango *k’eq-te* ‘blood’. This obstruent harmony scenario is more plausible than the other explanations proposed.

The contact-induced c’ reflexes of PHEC *t’ discussed above differ from the glottalized c’ arising through fusion in PNHEC (§3.3.1.4). The fusion process creates glottalized (secondary non-pulmonic) consonants from obstruent-glottal sequences rather than from inherited proto-phonemes or from donor languages. This distinction matters because fusion-derived c’ is regular and applies across the NHEC languages wherever the phonological conditions are met, whereas contact-induced c’ appears sporadically in specific lexemes.

PHEC *c’

After establishing the contact scenarios that explain irregular c’ reflexes of PHEC *t’ in certain HEC forms, the discussion now turns to examining lexemes that show regular c’ correspondences across all five HEC languages.

³⁵ As expected, most of the (Lowland) EC correspondences show t’/d’: [18] Had *t’iiga* ‘blood’, Oro *d’iiga* ‘blood’, Dir *diik* ‘blood, sap’, Mos *d’iika* ‘blood’, Kon *diika* ‘blood’, Ren *qliig* ‘blood’, Som *qiig* ‘blood’, Maay *qiig* ‘blood’, Boni *d’iik* ‘blood’, Arb *d’iik* ‘blood’, Dah *diiga* ‘blood’, cf. PWR **ts’eedee* ‘blood’.

³⁶ A broader hypothesis of “obstruent harmony” (between C₁ and C₂) might alternatively explain other instances of irregular correspondences that have been given various other interpretations in this study. For instance, the problematic glottalization of the labial: PNHEC [49] *k’uub- ‘egg’ in Kam *k’up’p’a-ta* ‘eggs’, Sid *k’uup’p’e* ‘eggs, egg’, Ged *k’uup’e*; *k’uup’p’e* ‘eggs, egg’, and the full assimilation in Had *k’uunk’a* ‘egg’. Obstruent harmony accommodates sibilant, velarization, glottalization, voicing, and other forms of assimilation (§2.5.2).

70 A Lexical and Phonological Reconstruction of Highland East Cushitic

These regular correspondences support the reasonable reconstruction of PHEC ***c'** as a distinct proto-phoneme.

PHEC ***c'** or ***tʃ** is a non-pulmonic post-alveolar affricate. The reconstruction of PEC ***c'** Ehret (1987) and Arvanites (1991, pp. 158–159) is supported by evidence across HEC languages and correspondences from East Cushitic, such as Sidaama **c'**, Oromo **c'**, Somali **j**, and Afar **d'**. Ehret (1987, p. 34) reconstructs **c'** for both PEC and PC, with consistent reflexes in Yaaku and Oromo **c'**. This section analyzes the regular and irregular reflexes of PHEC ***c'**, the EC correspondences, and potential external influences.

The reconstruction of PHEC ***c'** is based on regular correspondences in initial and non-initial positions and cognates found across the HEC languages. The examples below present the primary evidence that supports the reconstruction of PHEC ***c'** in initial position:

PHEC	[516] *c'il-	'to defecate'	[186] *c'ub-	'dip in, immerse'	[348] *c'iiʔ-	'bird'
Burji	c'il-	'to defecate'	c'uub-	'dip in, immerse'	c'iid'aa	'bird'
PNHE	*c'il-	'to defecate'	*c'ub-	'dip in, immerse'	*c'iiʔ-i-ʃʃ-	'bird'
C						
Had	c'ir-	'to defecate'	c'ufc'ufa	'dip in, immerse'	c'iiʔ-cco	'bird'
Kam	c'in-	'to defecate'	c'uf aʔ	'dip in, immerse'	c'ii-ccu-	'bird'
Sid	c'il-	'to defecate'	c'uuʔi	'dip in, immerse'	c'eʔ-icco,	'bird'
Ged	c'il-	'to defecate'	c'uup'-	'dip (in)'	c'iʔa	'bird'
cf.	-	-	c'uup'-	'dip,	-	-
Oro				baptize'		

In **[186]**, the change of ***b** to **f** in Hadiyya and Kambaata and **p** in Gedeo is explained via morpheme boundary alternations in §3.3.1.1. There is a subsequent reduplication of the same root in Hadiyya. The **uu** vowel doubling in Burji and Sidaama is likely due to contact with Oromo or internal alternation processes. In set **[516]**, the **n** in Kambaata **[516]** *c'in-* 'to defecate' is irregular.

For additional examples, see PHEC **[516]** **c'il-* 'to defecate' with **[517]** **c'iil-* 'excrement', **[186]** **c'ub-* 'dip in, immerse', with **[286]** **c'ob-ʔ-* 'to drip, to leak' and **[447]** **c'ubb-* 'sin', **[348]** **c'iiʔ-* 'bird' with **[226]** **c'iwʔ-*, **c'awʔ-* 'chick(en)'.

There are three more examples of initial **c'* reconstructable to PNHEC: [21] **c'ak'as-* 'to listen, spy', [416] **c'in-*³⁷ 'to squeeze', and [44] **c'a?* 'to clean'.

In non-initial position, additional evidence for PHEC **c'* is found in stems that retain regular reflexes as follows:

	[580]		[491]		[807]	
PHEC	<i>*boc'</i> -	'to carve, split'	<i>*fac'</i> -	'to cut, break'	-	-
Burji	<i>boc'</i> -	'carve'	<i>fac'</i> -	'to cut'	-	-
PNHEC	<i>*boc'</i> -	'to carve, split'	-	-	<i>*inc'</i> -	'to chew, twist'
C						
Had	<i>boc'</i> -	'to carve, split'	<i>fac'ite</i>	'burst of swelling'	<i>inc'iic'</i> -	'twist'
Kam	<i>boc'(c')-</i>	'snap (to break)'	-	-	<i>inc'iic'</i>	'twist'
Sid	<i>boc'</i> -	'to carve, split'	<i>fac'i</i>	'break a branch'	<i>c'-</i>	
Ged	<i>boc'</i> -	'cut, chop'	-	-	<i>hiinc'i</i>	'chew'
					<i>inc'a</i>	'chew'
cf. Oro	-	-	-	-	<i>(h)anc'-</i>	'chewing gum'

All three of the above sets support the reconstruction of **c'* in non-initial consonant in PHEC, with reflexes in Burji and PNHEC exhibiting regular correspondences. Set [580] has the Burji parallel *annif-* 'to chew', whose status as a cognate is doubtful because of deviant semantics, but it could be argued to be an older form with *c'* in PNHEC resulting from *j* and *ʔ* fusion as discussed in §3.3.1.4. Other examples of non-initial HEC **c'* include PHEC [377] **c'anc'-* 'to make noise' and PHEC [162] **fic'c'-* 'to scatter'. Interestingly, most of the above non-initial correspondences exhibit the sequence *nc'*, including an epenthetic *-n-* (§3.4.3).

Many examples of PHEC **c'* align with EC parallels, providing wider supporting evidence for PHEC **c'*:

PHEC [186] **c'ub-* 'dip in, immerse': Oro *c'uup'a* 'dip, baptize', *c'up'-*, *c'ub-* 'dip in', Dir *c'ap-* 'be wet', Mos *c'abiso*, *c'ubolayca* '(be) wet, be corrupt', Kon *ʃopɕ* 'to dip into water', Arb *c'um-* 'wade', Ts'am *c'ab-* 'to rot, to become wet'.

³⁷ See Som *dijji* 'milk a little, squeeze a trickle from', *diiq* 'trickle (of milk)'.

72 A Lexical and Phonological Reconstruction of Highland East Cushitic

PHEC [226] **c'iwʔ-*, **c'awʔ-* 'chick(en)': → Oro *c'uc'-oo*, *c'uc'-ii*, *wic'-oo*, *wic'ii* 'chick', Dir *c'aac'ut-ét*, *c'uuc'-ét* 'chick, (small) chicken', Kon *ʃaaff-eeta* 'bird sp.', Bay *c'uc'ute* 'chicken', Boni *c'eeʔúú* 'bird sp.', Gaw *č'uč'uw-e* 'chicks'.

PNHEC [416] **c'in-* 'to squeeze': Oro *c'unfa* 'squeeze'.

PNHEC [44] **c'aʔ* 'to clean': Dir *c'ik'* 'wash (tr), clean from a curse of witchdoctor', Mos *c'ik'* 'wash', Kon *ʃaɕ-* 'wash'.

PHEC [84] **c'ab-* 'to cut, sharpen': Oro *c'ab-a* 'break, be broken', Ren *jebic*, *jébi* 'break (something)', Som *jebi* 'break', Maay *j'ibiyow*, *j'abow*, *j'ab-* 'break', Boni *c'ibiyy-*, *j'ibiyy-/c'iyy-* 'to break' (cf. PSC **tʷuf-* 'to trickle' and PC **c'abk'-* 'to break up, break apart' (tr) (Ehret, 1987, p. 40)).

PNHEC [21] **c'ak'as-* 'to listen, spy': Arb *tʃ'ak'as-* 'listen' 'hear'.

PHEC [377] **c'anc'-* 'to make noise': Oro *c'ac'c'isa* 'fight back, snap back', Dir *c'aac'c'-* 'discuss (noisily)', Mos *c'ac'* 'quarrel', Kon *ʃaaff-* 'make a high pitched cry', Som *jaar-i* 'difference, contention (in a contest)'.

PHEC [580] **boc'-* 'to carve, split (wood)': Arb *boʃ-* 'beat, hew', Bay *boc'i* 'beat, hew', Yaaku *pilc'* 'small sticks of firewood' (cf. Bilin *bec'bec'-*, *bet'bet'-* 'shake' (Lamberti & Sottile, 1997, p. 314), Som [321] *bujj* 'pierce', Maay *buj'iyow* 'pierce', Boni *bac'-iy-*, *baj'-iy* 'to pierce'.

PHEC [491] **fac'-* 'to cut, break': Oro *fac'c'ee* 'fighting with a stick or sword'.

PNHEC [807] **inc'-* 'to chew, twist': Oro *(h)anc'-aa* 'chewing gum', Som *ħanj-o* 'chewing gum', Ren *ħánja* 'gum from certain trees', Mos *ʔanc'uhiso* 'chew'.

The PNHEC stem **c'aac'urr-*, **c'iic'oorr-* 'chick(en)' with **rr** in NHEC formally aligns with the following EC parallels Oro *c'irrii* 'bird', Gaw *c'irraʕ-itt-e* 'red-billed oxpecker', Ts'am *c'uruq'e* 'bird sp.', and Arb *tʃ'aaraté* 'oxpecker'. However, those parallels semantically align with PHEC [348] **c'iiʔ-* 'bird'. Arbore probably has two archaic stems *tʃ'ak'as-* 'listen' and *d'eg(a)y-* 'hear', but one parallel in PNHEC [21] **c'ak'as-* 'to listen, spy'. According to Arvanites (1991, p. 172), in Somali [377] *jaar-i* 'difference, contention (in a contest)', **r** is

from *d'. Another root discussed above, PHEC [516] *c'il- 'to defecate', has no EC or areal parallel.

The number of sets with regular correspondences in both initial and non-initial positions in HEC, combined with the numerous EC parallels, supports the reconstruction of PHEC *c'.

Three potential contact sources could explain irregular c' in HEC: Oromo borrowing, Omotic substrate influence, and Amharic loanwords. Each requires a separate evaluation.

Despite their regular correspondence of *c' within PHEC, some of the above forms with Oromo parallels might nevertheless represent loans into HEC languages based on form and semantics, such as [186] or irregularity of correspondences with the rest of HEC, as in [162].

	Oromo	→ Burji	→ Ged	→ Sid
[186]	c'uup' - 'dip, baptize'	-	c'uup' - 'dip (in)'	-
[377]	c'ac'c'isa 'fight back, snap back'	c'anc' - 'to rant'	c'anc' - 'shout, cry'	c'anc' - 'noise, shouting'
[169]	k'unc'ee 'bark'	k'unc'e 'bark'	k'unc'e 'bark'	-
[162]	fac'c'aʔa 'be scattered, sown', fac'c'aasa 'scatter, sow'	fac'c' - 'scatter'	fac'aasa 'scatter'	-

The forms with cognates within the rest of HEC and EC might represent cases where the existing cognate has been replaced or include double reflexes as a result of the loans. See [169], [367], and [476] for more examples.

Contrary to the retained *c' proposed by Arvanites (1991) and Ehret (1987), Sasse (1979, p. 30) argues that c' in the HEC languages predominantly appears in loanwords, with two or three autochthonous instances that lack cognates outside HEC. Sasse's statement overgeneralizes and only aligns with instances where c' is irregular across HEC languages, which may reflect contact influences. The remaining paragraphs of this section review such irregular correspondences in the above examples with their EC parallels. Notable cases include:

74 A Lexical and Phonological Reconstruction of Highland East Cushitic

PHEC [186] **c'ub*- 'dip in, immerse': Afar *d'uum*- 'sink', Som *ku qumbi* 'plunge something into water, immerse', Gaw *k'ap*- 'to rot (of dead animals)' (cf. Dah *t'up*- 'to leak').

PHEC [226] **c'iwʔ*-, **c'awʔ*- and PNHEC **c'aac'urr*-, **c'iic'oorr*- 'chick(en)': Mos *cu-cute* 'chick', Som *qawqawle* 'bird sp., woodpecker'.

PNHEC [416] **c'in*- 'to squeeze': Som *qijji* 'milk a little, squeeze a trickle from' and *diiq* 'trickle (of milk)', Arb *d'angalay*- 'gush, pour out'.

PNHEC [44] **c'aʔ* 'to clean': Ren *qixa* 'wash (something)', Som *daq*- 'wash', Maay *qiyow*, *qiq* 'to wash', Boni *d'ak'aʔ*, *d'aʔ*- 'to wash', Arb *d'ik*- 'wash', Das *fik*, *gik*- 'to wash', Elm *d'ik* 'to wash', Yaaku *t'ok*- 'wash'.

PHEC [84] **c'ab*- 'to cut, sharpen': Dir *k'ep*- 'break', Mos *k'eḅ*- 'break', Kon *k'ep*- 'break', Bay *ebis*- 'break' (tr), Arb *k'eb*- 'break', Das *géb* 'to break'.

PNHEC [21] **c'ak'as*- 'to listen, spy': Oro *d'agaya* 'hear', Dir *ḏakay*- 'hear', Mos *ḏakay*, *ḏak'a* 'listen, hear', Kon *ḏak-ay*- 'hear', Saho *-eqeg*- 'know, recognize', Afar *qag* 'auricular region', Ren *qaaga* 'hear', Som *qeqeyso* 'hear, listen to', Maay *qeqyow* 'hear, listen', Boni *dijésio* 'listen', Elm *d'ekait*, *d'ey* 'to hear', Yaaku *dek*- 'hear'.

PHEC [580] **boc*'- 'to carve, split (wood)': Oro *bat'bat* 'shake', Bilin *bec'bec*'-, *bet'bet*'- 'shake' (Lamberti & Sottile, 1997, p. 314).

PHEC [491] **fac*'- 'to cut, break': Mos *faats'o* 'branch (of tree)'.

For PHEC [186] **c'ub*- 'dip in, immerse', Sasse (1979, p. 30) reconstructs PEC **d'ub*- 'dip in'. Parallels of PHEC [226] **c'iwʔ*-, **c'awʔ*- and PNHEC **c'aac'urr*-, **c'iic'oorr*- 'chick(en)' are attested in Saho as a loanword *s'us'uut* 'chicken' from Tigrinya/Arabic, as claimed by Esayas Tajebe (2015, p. 30), with no matches in Omotic. It is also likely an onomatopoetic root. In [21], Arbore shows both **d'** and **c'** variants, potentially reflecting a distinction between two separate roots: one with **c'** (← ***c'**) and one with **d'** (← ***t'**). If other EC languages have merged these roots, the two reflexes here would represent regular developments from their respective proto-forms. Sasse (1979, p. 17) reconstructs PEC **d'eg*-/**dog*- 'to hear' (cf. PC **dleg*- 'to listen' (Ehret, 1987, p. 83)). In [580] **boc*'- 'to carve, split (wood)', Oro *bat'bat* 'shake', has been transferred into Gedeo as *bat'a* 'split open, spread'. This root has parallels in Bilin *bec'bec*'- and *bet'bet*'- 'shake' (Lamberti & Sottile, 1997, p.

314). The existence of both **c'** and **t'** in the parallels complicates the cognacy relationship.

The most common correspondences in the above examples are **d'** with a few cases of **k'**/**g**, **d**, and a single attestation of **c**, **f**, **ø**, **ts'**. Establishing a regular palatalization rule to explain **t'/d' → c'** is difficult,³⁸ but as Arvanites (1991, p. 130) proposes, it might be an areal change.³⁹

Several sets show regular **c'** correspondences across all HEC languages, but multiple factors suggest they are Amharic loanwords: (1) limited or absent EC cognates outside HEC, (2) cultural domains (clothing, weapons, stimulants), and (3) close phonological matches with Amharic forms.

Amharic		P(N)HEC		→ HEC
<i>c'ork'a</i> 'rag'	[220]	* <i>c'ark'</i> -	'rags, cloth'	Burji <i>c'ark'aa</i> , Ged <i>c'ark'e</i> , Had, Kam <i>c'ark'a</i>
<i>c'ube</i> 'dagger'	[544]	* <i>c'ub</i> -	'dagger, knife'	Had <i>c'ube?</i> e, Kam <i>c'up'i-ta</i> , Ged <i>c'uup'e</i> , Sid <i>c'uube</i>
<i>c'amma</i> 'shoe'	[81]	* <i>c'amm</i> -	'shoes' ⁴⁰	Ged <i>c'aamma</i> , Sid <i>c'amma</i> , Had <i>c'aamma?</i> a, Kam <i>c'amma</i>
<i>c'at</i> 'chat (<i>Catha edulis</i>)'	[513]	* <i>c'aat</i> -	'chat (<i>Catha edulis</i>)'	Had, Kam <i>c'aata</i> , Sid, Ged <i>c'aate</i>
<i>c'äbt'</i> 'gonorrhea'	[490]	* <i>c'obt'</i> -	'gonorrhea'	Burji <i>c'ooddoo</i> , Had <i>c'abt'o?</i> o, Kam <i>c'obt'a</i>

Despite the complicating factor of Amharic loanwords showing regular **c'** correspondence, the overall pattern for PHEC (i.e., inherited) ***c'** is that it has regular correspondences in HEC, it has more (and mostly regular) parallels in (East) Cushitic, and the reconstructed words have fewer parallels in the neighboring non-Cushitic languages. The attested Oromo loans appear to involve double reflexes, i.e., reflexes of original PHEC forms alongside borrowed reflexes in the cognate HEC and Oromo lexemes.

³⁸ Is it possible that, in those cases, ***c'** irregularly changed to **k'**?

³⁹ According to Arvanites (1991, p. 130, 157), "[n]ot to be overlooked is the fact that ***t'** sometimes is perplexing!; Correlatable with ***k'**. The way that this is accomplished via regular phonetic change involves palatalization (a common areal change involving alveolars) of ***t'** to **c'**."

⁴⁰ Reconstruction [81] actually is **kob-?* 'shoe, sandal' but **c'amm-* is proposed as an additional provisional reconstruction. Amharic *c'amma* 'shoe' is from Lamberti & Sottile (1997, p. 331).

The reconstruction of PHEC ***c'** thus rests on regular correspondences across all five HEC languages, supported by wider EC parallels. This non-pulmonic ***c'** must be distinguished from glottalized ***c'** arising through the fusion of sibilants with the glottal stop in NHEC. The fusion process (§3.3.1.4) creates **c'c'** geminates wherever **ʃ** or **s** precedes **ʔ**, yielding forms that superficially resemble inherited ***c'** reflexes but follow a different developmental path.

Summary

The reasonable reconstruction of PHEC ***t'** and ***c'** as distinct proto-phonemes rests on systematic correspondence patterns across all five HEC languages, supported by extensive attestation in both word-initial and non-initial positions. Each phoneme shows regular reflexes: ***t'** yields **t'**, while ***c'** consistently yields **c'** across HEC. The complexity arises not from the robust regular correspondences themselves but from explaining why subsets of lexemes show irregular reflexes.

The irregular **c'** and **d'** reflexes in the ***t'** correspondence set reflect multiple contact scenarios rather than undermining the reconstruction. First, direct borrowing from Oromo (which has **d'**) introduces **d'** into specific HEC lexemes, particularly in Burji, Gedeo, and Sidaama, due to sustained contact. Second, analogy and harmony processes operating within HEC can produce secondary **c'** from inherited ***t'** in specific phonological environments. Third, earlier contact with Omotic languages contributed additional irregular patterns with **c'** or **t'**. Crucially, these contact-induced irregularities are identifiable precisely because the regular inherited patterns are so clear.

The reconstruction of ***c'** as distinct from ***t'** receives further support from three independent lines of evidence: (1) numerous EC cognates showing systematic ***c'** correspondences; (2) the semantic distribution, wherein ***c'** reconstructions cluster in basic vocabulary unlikely to be wholly borrowed; (3) the phonological naturalness of the contrast, which exists throughout Cushitic and reflects the articulatory distinction between apical (**t'**) and post-alveolar (**c'**) non-pulmonics.

This analysis demonstrates the methodological approach outlined in §1.5: treating contact and inheritance as a continuum rather than binary alternatives. Some lexemes inherit ***t'** regularly, others inherit ***c'** regularly, and still others show complex histories involving borrowing, analogy, or secondary processes. This nuanced approach provides a more historically realistic reconstruction than forcing all data into rigid categories.

2.1.3 Labials

Highland East Cushitic has four labial consonants: the stop ***b**, the nasal ***m**, the fricative ***f**, and the glide ***w**. This section examines the ***b**, ***m**, and ***f**, while the labial glide ***w** is discussed with the other glide ***y** in Section 2.1.4. Where ***b** lenition involves an intermediate ***w** stage, only the necessary relevant evidence is presented here; the full discussion of ***w** reflexes follows in the PHEC ***w** section. Their historical development reflects three key processes that shaped their modern reflexes across the daughter languages:

- a. Spirantization of intervocalic ***b** to [β] in Burji, Sidaama, and Gedeo
- b. Lenition of intervocalic ***b** to ***w**, sometimes with subsequent loss
- c. Loss of intervocalic ***w**

These three processes form a feeding relationship, highlighting the historical tendency to weaken labials (***b** and ***w**) in intervocalic position. While the fricativization is allophonic, the lenition of ***b** appears to be often triggered by more recent boundary alternation processes compared to that of ***w**, though they may have operated simultaneously at an earlier stage.

PHEC ***b**

PHEC ***b** is securely reconstructed based on evidence in both initial and non-initial positions.

PHEC	[34] *baʔ-	'be lost'	[9] *arrab-	'tongue'	[183] *gib-	'to refuse, hate'
Burji	be(y)-, bay-	'be lost'	arraaba	'tongue'	jibb-	'to hate'
PNHEC	*baʔ-	'be lost'	*arrab-	'tongue'	*gib-	'to refuse'
Had	biʔ-	'to be lost'	allabo	'tongue'	gib-	'flee'
Kam	baʔ-	'be lost'	arrabi-ta	'tongue'	gib-	'refuse'
Sid	baʔi	'perish'	ʔarrabo, arrawo	'tongue'	gib-	'refuse'
Ged	baʔ-	'be lost'	arrabo	'tongue'	gib-	'refuse'

In [183], the geminate **bb** in Burji might be due to the influence of Oromo *jibba* 'dislike'.⁴¹

⁴¹ "The Burji and Gedeo words could be purely borrowings from Oromo [jibba], but perhaps the Oromo word, presumably a distant cognate, has just influenced these forms" (Hudson, 1989, p. 7). Additionally, in Sidaama, "assimilation takes place when

Additional examples include PHEC [444] **bagaz-* ‘spear’, PNHEC [460] **beet-* ‘son, child’, PHEC [273] **biin-* ‘mosquito, fly (sp.)’ (see Amh *bimbi* ‘mosquito’), PHEC [407] **ebel-*, **ebal-* ‘so-and-so’ (likely compound word), PNHEC [112] **bub-* (intr) ‘burn’.

Burji and Gedeo show spirantized variants **b** ~ **β** intervocalically (Sasse, 1982, p. 16; Eyob Kelemework, 2015, p. 34), while Sidaama exhibits both **β** and **w** variants of **b** intervocalically (Yri & Pepper, 2019; Anbessa Teferra, 2000, p. 26; Hudson, 1989, p. 12; Bender, 2020, pp. 116–117). This variation is reflected in different transcription practices across sources: Yri & Pepper (2019) use phonemic /**b**/ and orthographic <**w**>, Gasparini (1983) and Indrias Xa’miso et al. (2007) have [**w**].⁴² The data is cited in this study without modification in this regard. The spirantization **b* → [**β**] in these three languages represents an initial stage of lenition without complete loss when stem-final **b** is followed by a vowel through morphophonological processes. Examples from Sidaama include:

b ~ **β**: *ʔaaba* ~ *aaβa* ‘father’

b ~ **w**: *ʔarrabo* ~ *arrawo* ‘tongue’

duub- ‘be satisfied’, *duuwí* (perf) ‘He was satisfied’⁴³

The nouns are followed by the terminal vowels **-a** and **-o**, and the verb is followed by the perfective **-í**. In [252] *hasaaw-i* ~ *hasaab-i* ‘speak’, even an original **w** has an alternant **b**.

A distinct lenition process affecting intervocalic **b* predates the synchronic spirantization described above. This older rule resulted in the complete loss of **b* in intervocalic environments:

ii. Lenition of intervocalic **b*

b* → **w* / **V **V**

the suffix /-tú/, 3rd f. / 3rd pl. perfect marker, is attached to stem-final obstruents” such as in *gib-* ‘refuse’ *gibbú* (perf) ‘she/they refused’ (Anbessa Teferra, 2000, p. 21).

⁴² “It is rather unfortunate that the decision makers chose to write the allophone [**β**] of /**b**/ as <**w**> intervocalically rather than <**b**>, the more so as the intuitively correct <**b**> is maintained word initially and <**bb**> for the geminated one word medially. The phoneme /**w**/ occurs word initially and in geminated form it contrasts with /**bb**/ word medially as /**ww**/. /**w**/ has lip rounding in all positions while [**β**] (allophone of /**b**/) is only rounded next to vowels with lip rounding, viz. /**o**/ and /**u**/. And this environmentally induced lip rounding is probably the cause of the unfortunate choice” (Yri & Pepper, 2019).

⁴³ Data from Anbessa Teferra (2000, p. 26) and Hudson (1989, p. 7).

Evidence for this historical process can be found in Burji, which retains the ***b** in consonant clusters, such as **bd**, while NHEC languages show a complete loss of this sound in intervocalic positions. The lost ***b** can be reconstructed based on three lines of evidence: Burji cognates, East Cushitic cognates, and NHEC vowel correspondences.

PHEC [115] **k'ab y-t-* 'become cold' → Burji *kabb-* 'cold, cool, become cold'; PNHEC **k'iid-* 'become cold' → Had, Sid, Ged *k'iid-* 'cool, become cold', Kam *gid-* (cf. Oro *k'abb-an-* 'become cool (liquid)').

Beyond the assimilation in Burji of ***bd** to **bb**, the corresponding vowel patterns and the final consonant in NHEC support the reconstruction of the verbalizing agentive denominative ***y-t** (§4.2.5.3). The ***b** in this root is additionally attested in other EC cognates (PEC **k'ab-*), with the final **d** likely deriving from the fossilized ***-t** in HEC. The loss of ***b** in NHEC can be reconstructed in three phases:

- (a) Creation of an intervocalic environment via the verbalizer vowel **i** (§4.2.5.2): **k'ab-i-d-*
- (b) Lenition of ***b** to **w**: **k'awi-d-*
- (c) Loss of intervocalic ***w** and merger of ***a** and ***i** into long **ii**, yielding PNHEC **k'iid-*

While this reconstruction accounts for most of the developments, the initial velars in Burji (**k**) and Kambaata (**g**) remain problematic in light of their correspondence with non-pulmonic **k'** in the rest of HEC.

Hudson (1989, p. 12) notes that there is no evidence for reconstructable intervocalic ***b** within roots in HEC. Synchronically, intervocalic **b** appears only in borrowed words and a few retained exceptions. This is exemplified by [349] **gabal-* 'border, limit', which shows both irregular intervocalic **b** in all HEC and irregular **l** in Hadiyya (where **r** would be expected §2.1.1.3). The exceptions include PHEC [214] **t'aban ʔ*, **t'aban ʔ-s-* 'to slap' (with loss in Sidaama *d'aʔa assa* 'to slap (the face)' only via lenition ***ø** ← ***w**) and PHEC [606] **gubeed-* 'thigh' (see Oro *gudeeda*). These represent arguable cases of exceptional retention, where morphological boundaries or lexical factors may have blocked the otherwise regular intervocalic lenition process. The limited number of such exceptions (compared to the overwhelming pattern of loss) strengthens rather than undermines the reconstruction of the regular rule.

80 A Lexical and Phonological Reconstruction of Highland East Cushitic

Non-initial ***b** is not affected by the lenition rule when it occurs in a geminate or in a cluster, as demonstrated in the following cognate sets:

	[270]		[3]		[95]	
PHEC	*t'ibb-	'hundred'	*aabb-	'father'	jomb-	'lung'
Burji	c'ibba	'hundred'	aabboo	'father'	jombi	'lung'
PNHEC	*t'ibb-	'hundred'	*aabb-	'father'	-	-
Had	t'ibbe	'hundred'	aabba	'father'	-	-
Kam	t'ibbi-ta	'hundred'	aabba	'father'	-	-
Sid	t'ibbe	'hundred'	?aaba	'father'	jombo	'lung'
Ged	(d'ibba)	'hundred'	abbo	'uncle'	jombo	'lung'

More complex correspondences emerge in examples with vowel alternations and ***lb** clusters.

PHEC **[4]** **gulib-*, **gilub-* 'knee' → Burji *gilba* 'knee'; PNHEC **gulub-* → Had *gurubbo*, Kam *gulubi-ta*, Sid *gulucco*, Ged *gulubo* (cf. Oro *jilba* 'knee', PEC **gilb* → **gilib* → **gilub*).

The correspondence patterns and vowel reflexes require careful analysis. While rounding of V_2 (***i** to **u** before **b**) followed by V_1 assimilation might explain the NHEC forms, the comparative evidence points to a different reconstruction. Three independent lines of evidence converge on **gulib-* as the PHEC form: (1) the lack of palatalization in the Burji form indicates V_1 was already ***u** at the proto-level, (2) the Oromo cognate *jilba* preserves the expected ***i** in V_1 position before ***l** (consistent with PEC **gilb-*), and (3) the NHEC vowel assimilation (**gulib-* → **gulub-*) follows attested patterns of labial-conditioned rounding. This reconstruction demonstrates the application of the comparative method, where multiple phonological processes operating at different chronological stages must be disentangled to arrive at the most parsimonious proto-form.

Some HEC languages exhibit problematic correspondences involving the sounds **b**, **m**, and **mb**. The nasal variants **m/mb** appear specifically in Burji, Gedeo, and Sidaama examples:

[514] Burji *amb-* ‘bring’: PHEC **ebb-* ‘to bring’ → PNHEC **ebb-* ‘to bring’
→ Had and Kam *eeb-* ‘bring, marry (a woman)’, Sid *ʔabb-*
‘bring’, Ged *eʔijfa* (← **ew-ij-*)⁴⁴ ‘bring in’.

[495] Gedeo *lumo* ‘big’: PHEC **lab-* ‘many, big, much’ → Burji *laboo*,
Had *lob*, Sid *lobo*.

[550] Sidaama *ʔibiicco* (f) (Yri & Pepper, 2019); *imiawe* (sg), *iwiawe* (pl)
(Indrias Xa’miso et al., 2007, p. 273); *ibi-cco*, *ibiibe* (Hudson,
1989, p. 94) ‘louse, lice’: PHEC **ibb-* ‘louse’ → Burji *ibba*;
PNHEC **ib-*, **ib~ib-* ‘louse’ → Had *ibiiba*, Kam *ibiibi-ta*; *ibiibe*
‘louse’, Ged *ibbe*.

These irregular correspondences can be explained through epenthetic **-n-** (§3.4.2), as evidenced by cognates in other HEC and EC languages (e.g., [514] Jiddu *eeb-* ‘bing’ and Das *ʔeeb-* ‘bring’). The following development pathway can be proposed:

[514] Burji <i>amb-</i>	‘bring’	← <i>*a-n-b-</i>
[495] Gedeo <i>lumo</i>	‘big’	← <i>*lu-n-b-</i>
[550] Sidaama <i>imiawe</i>	‘lice’	← <i>*i-n-biib-</i>

Based on this interpretation, **m** represents an inserted and assimilated **-n-** before the labial, found as the cluster **mb** in [514] but reduced to **m** in [495] and [550].

In summary, PHEC ***b** shows the following key patterns: preservation in initial position, occurrence in non-initial position primarily as geminates (some potentially resulting from old assimilation of ***b** and another consonant (§3.1.1)) or in clusters, loss of intervocalic single ***b** via the glide ***w** following a pattern common to other EC languages, and glottalization and occasional fricativization in final position when followed by voiceless consonants (**ʔ**, **s**, **j**, or **c**) in morpheme boundary alternation processes (§3.3.1).

PHEC ***m**

The labial nasal ***m** is securely reconstructed with evidence from regular reflexes across the HEC languages. Unlike the alveolar nasal ***n** (see

⁴⁴ Comparable with the ***b** loss in Sid *d’aʔa assa* ‘to slap (the face)’ ← PNHEC [214] **t’aban ʔ-s-* ‘to slap’.

82 A Lexical and Phonological Reconstruction of Highland East Cushitic

§2.3.1.4), ***m** occurs regularly in both initial and non-initial positions, as demonstrated in the following cognate sets:

PHEC	[54] *min-	'house'	[72] *lam-	'two'	[48] *tum-	'to pound'
Burji	mina	'house'	lám-a	'two'	tum-	'to hit/churn'
PNHEC	*min-	'house'	*lam-	'two'	*tum-	'to pound'
Had	mine	'house'	lamo	'two'	tum-	'to hit/strike'
Kam	mini	'house'	lamu	'two'	tum-	'to hit/strike'
Sid	mine	'house'	lame	'two'	tum-	'to hit/strike'
Ged	mine	'house'	lame	'two'	tum-	'to pound'

Further examples include PHEC [\[68\]](#) **malañ-* 'pus', [\[268\]](#) **mur-* 'to cut', [\[547\]](#) **am-* 'mother', [\[675\]](#) **k'oom-* 'calabash', and [\[391\]](#) **kum-* 'thousand'.

While ***m** shows general stability, it undergoes important morpheme boundary alternation processes in specific environments. Most notably, stem-final ***m** exhibits partial assimilation to **n** before alveolar-initial suffixes (see §3.1.1.4 for a detailed discussion). This process remains productive in modern HEC languages, as evidenced by Sidaama's synchronic assimilation of stem-final **m** to **n** before **t** (Kawachi, 2007, pp. 49–50).

PHEC ***m** also plays a crucial role in the morphological system in the form of the reconstructable passive suffix **-am** (see §4.2.1). The presence of the labial nasal **m** in this morpheme is consistent across HEC, along with its regular phonological alternations, which supports its reconstruction for PHEC.

PHEC ***f**

The reconstruction of the labial fricative ***f** for PHEC is secure and supported by regular initial and non-initial reflexes across all HEC languages. The following cognate sets demonstrate these stable correspondences:

PHEC	[167] *fayy-	'healthy'	[70] *afoʔ-	'mouth'	[40] *tuf-	'to spit'
Burji	fayya-n-ga	'healthy'	afeyi	'mouth'	tuf-	'to spit'
PNHEC	*fayy-	'healthy'	*afoʔ-	'mouth'	*tuf-	'to spit'
Had	fayyaʔa	'healthy'	afoʔo	'hole'	tuf-	'to spit'
Kam	fayya	'healthy'	afoo	'mouth'	tuf-	'to spit'
Sid	fayyaaleessa	'healthy'	ʔafo	'mouth'	tuf-	'to spit'
Ged	fayya	'healthy'	afaʔo	'mouth'	tuf-	'to spit'

Further evidence for PHEC *f comes from PHEC [339] *ful- ‘to go out, exit’, [415] *afal- ‘liver’, [47] *k’uf- ‘to cough’, and PNHEC [65] *faras- ‘horse’.

Unlike the labial stop *b, there are no reconstructable instances of geminate *ff in HEC. The few cases of ff that do appear are clearly onomatopoeic in nature (cf. Kambaata *uff aʔ-* and Sidaama *uffu ass-* ‘blow (on fire)’).

There is limited evidence for PHEC *f. The majority of the instances of the fricative f are found in loanwords in the HEC languages, and it appears more frequently in Burji. There are instances of non-initial PHEC *f developing from an earlier *b discussed in §3.3.1.1. Another change *s → f showing Oromo influence on Burji is discussed in §2.3.1.2, PHEC *s.

Summary

The labial consonant series (*b, *m, *f) is securely reconstructed for PHEC, with systematic correspondences across all five HEC languages supporting each proto-phoneme. The examination of labial consonants in HEC reveals systematic patterns in both their historical development and synchronic distribution. While there is remarkable stability across all labials in initial position, non-initial environments trigger various developments, most notably in *b, which undergoes spirantization, lenition, and eventual loss in intervocalic position. The nasal *m demonstrates greater stability, showing only limited alternations through partial assimilation before alveolar-initial suffixes, while also serving as a crucial morphological marker in the passive suffix -am. The fricative *f, though it does occur in geminate form, exhibits regular correspondences across the languages, with some instances developing from earlier *b through conditioned changes at morpheme boundary. These developments align with broader tendencies in Cushitic historical phonology, particularly in the preference for lenition in intervocalic positions and the stability of initial consonants, while also revealing innovations specific to the HEC languages.

2.1.4 Glides

The glides *w and *y pose distinctive reconstruction challenges. While their reflexes in initial position show regular correspondences, intervocalic positions show extensive loss, making it difficult to distinguish inherited glides from epenthetic segments that emerge to resolve hiatus (§§3.4.1.2, 3.3.4).⁴⁵

⁴⁵ Both *w and *y serve important morphophonological functions in the HEC languages, appearing in derivational processes (§4.1.1.2, §4.1.5.2, §4.2.5, §4.2.4.1).

This section examines the historical development of these glides and their distinctive patterning across the HEC languages, with particular attention to positional constraints and their interaction with surrounding vowels.

Glide rules

- (a) Loss of intervocalic glides
- (b) Gemination rule of intervocalic glides

PHEC *w

The following sets show the reflexes of the securely reconstructed PHEC *w:

PHEC	^[293] *waʔ-	'water'	^[304] *wit'-	'to sow, scatter'	^[252] *hasaaw-	'speak'
Burji	waa	'water'	wat'ar-	'scatter, spread'	hasaaw-	'to talk'
PNHEC	*waʔ-	'water'	*wit'-	'to sow, scatter'	*hasaaw-	'speak'
Had	woʔo	'water'	wit'-	'to sow, to seed'	hasaaw-	'speak'
Kam	woʔa	'water'	wit'-	'sow, scatter'	haasaaww-	'discuss'
Sid	wa	'water'	wit'-	'sow'	hasaaw-	'speak'
Ged	wedeʔ-	'water'	wiisa	'sow'	haasoʔ	'speak'

In ^[293], rounding of *a → o after w is plausible in Hadiyya and Kambaata. However, all Gedeo examples are problematic: ^[293] *wedeʔ-e* and ^[304] *wiisa* have unique **d** and **s**, respectively,⁴⁶ while ^[252] *haasoʔ-* has lost the non-initial **w** altogether. Other reconstructable series with non-initial *w are difficult to find, and none exist for *w in intervocalic position.

More examples of PHEC *w include PHEC ^[139] *wozan- → PNHEC *wozan- 'heart', PHEC ^[259] *wol- → PNHEC *wol- '(each) other', PHEC ^[309] *war- → PNHEC *war- 'to gossip, tell the news', PHEC ^[440] *weeʔn- → PNHEC *weeʔn-, *weeʔn-i-ff- (sgv) 'Colobus monkey', PNHEC ^[406] *wees- 'ensete tree (Ensete ventricosum)', and PHEC ^[252] *hasaaw- → PNHEC *hasaaw- 'speak, talk', and PHEC ^[443] *uħw- → PNHEC *uħw- 'to give'.⁴⁷

⁴⁶ See Sheko *yiits* 'to sprinkle'.

⁴⁷ See PEC *-ħiw-/*-ħuw- (or *-hiw-/*-huw-) → PHEC *ʔuhw- 'give' (Sasse, 1982, p. 185).

The intervocalic glide loss $*w \rightarrow \emptyset$ rule ((iii)a) may represent a further development of the lenition process $*b \rightarrow *w$ rule (ii), though the chronological relationship remains uncertain.

iii. Loss of intervocalic glides

- a. $*w \rightarrow \emptyset / V_V$
- b. $*y \rightarrow \emptyset / V_V$

The scarcity of evidence for non-initial $*w$ and the lack of any single intervocalic $*w$ indicate that this must be an old rule (attested in East-Omo-Tana as well).

There are limited instances of non-initial w in clusters and geminates.

w?: PHEC [226] $*c'iw?$ -, $*c'aw?$ - 'chick(en)' → Burji $c'iw?$ e; PNHEC $*c'aac'urr$ -, $*c'iic'oorr$ - via $*c'ac' \sim c'u$ - ($*c'iu$ -e) → Ged $c'aac'ute$, Sid $c'aac'c'urre$, and via $*c'ic' \sim c'o$ -rr- (pl) → Had $c'iic'oolla$, $c'iic'oora$, Kam $c'iic'oorra$.

***wm**: PHEC [399] $*dawm$ - 'red' → Burji $duwwaaga$ (m), $duwwaada$ (f); PNHEC $*dawam$ - 'red' → Had $daama?a$ 'dark red', Kam $daama$ 'specific colour of a cow (red and black, e.g. front black, back red)', Sid $duume$ 'red'.

***wh/*wk**: PHEC [466] $*diwh$ -, $*diwk$ - 'empty' → Burji $duwwa$ 'empty and deserted house'; PNHEC $*diwh$ -, $*diwk$ - 'empty' → Had $diiha$ 'empty', Kam $diiha$ 'empty', Ged $duwwi$ 'empty, bare, naked'.

The sequence development of the examples is as follows:

PHEC $*iw?/*aw?$, $*awm$, $*wh/*wk \rightarrow$	Burji $iw?$ -, uww , uww
	PNHEC $*ii/*aa$, $*awam$, $*iwh$
	→ Had and Kam ii , aa , ii
	Ged and Sid aa , Ged uww , Sid uu

In [399] and [466], Gedeo has $diimmo$ 'red' and $duwwi$, $duwwika$ 'empty, bare, naked' and Burji has $duwwa$ 'empty and deserted house', which may reflect influence from Oromo $diimaa$ 'red' and $duwwaa$ 'empty'.

Gemination of some single intervocalic glides (in Burji (Tesfaye Baye, 2015, p. 43), Sidaama, and Gedeo) is provisionally proposed. However, there are more instances of single ***y** gemination than ***w**.

***w** → **ww** /V_V (rare)

***y** → **yy** /V_V

The ***w** → **ww** /V_V is possibly marginal or morphophonological in some HEC languages, such as Sidaama, Gedeo, and Burji. It can phonotactically block the loss rule of single intervocalic ***w** (iii). That is mainly evident in the plurative suffix ***-uw-a**, which has been retained in Hadiyya, Sidaama, and Gedeo as **-u(w)wa** (see §4.1.1.2, Plurative ***-uw-a**).

Single intervocalic **w** is geminated in the Amharic word *billawa* ‘knife’ in Hadiyya and Kambaata [96] *billawwa*, but merges with **b** and is retained as single in Sidaama *billaba* (cf. Oromo, which has **ø** in *billaa* ‘knife’). Kambaata [147] *hawwaru-ta* ‘evening’ is from a weakened ***b** of PNHEC **habar* ‘evening’. Consequently, such **ww** geminates are not original (especially found in the plurative suffix ***-aw-a** → ***-aww-a**). Intervocalic gemination affects **y** more frequently than **w**, with numerous attested examples for **y** but limited evidence for **w**.

There are instances of Burji **h** corresponding to NHEC **w** in initial positions, e.g., Burji [520] *huuw?* ‘to fill, to fill up’ and Sidaama *wo?m-* ‘all, every, full’, with the **w** as a reflex of ***h** before ***u** in NHEC.

PNHEC ***y**

Unlike the PHEC glide ***w**, the examples of initial ***y** are limited.

	[388]		[28]		[284]	
PHEC	*y	‘to say, do’	*ay-	‘who?’	*ayyaan-	‘festival’
Burji	iy-, i-	‘say, do’	ayye	‘who?’	ayyaanaa	‘festival’
PHEC	*y	‘to say’	*ay-	‘who?’	*ayyaan-	‘festival’
Had	y-	‘say’	ayye	‘who?’	ayyaana	‘festival’
Kam	y-	‘say’	aye	‘who?’	ayyaana	‘festival’
Sid	y-	‘say, judge’	aye	‘who?’	?ayyaana	‘festival’
Ged	hiyy-	‘say, try, do’	ayye	‘who?’	ayyaana	‘festival’

PHEC [388] ***y** ‘to say, do’ becomes Gedeo *hiyy-* with initial **h-** and is similar to Burji where **-i** insertion has occurred. Section 4.2.4.1 includes a detailed

discussion on this (support) verb within HEC.⁴⁸ In PHEC [28] **ay-* ‘who?’, the geminated **yy** of Burji, Gedeo, and Hadiyya is likely due to an epenthetic **-y-**, while Sidaama and Kambaata *aye* (← **a-y-e*) have inserted **-y-** after the loss of intervocalic single ***y** via rule (ii(b)), all after the suffixation of the vocalic morpheme.⁴⁹ Eyob Kelemework (2015, p. 107) states that “[t]he nominative form of the pronoun ‘who’ is [ayye]; the degeminated form [aye] and the short form *aa* are also used as an interrogative pronoun” in Gedeo. There are similar cognates with geminate **yy** and single **y** in the rest of EC. Finally, PHEC [284] **ayyaan-* ‘festival, holiday’ has Cushitic cognates but is possibly a loanword. It has parallels in Ethiosemitic (Sasse, 1982, p. 30).

Further evidence for PHEC ***y** includes PHEC [749] **yaaʔ-* ‘assembly, meeting’, [167] **fayy-* ‘healthy’, [91] **mooyy-* ‘mortar’, and [448] **gaayy-* ‘pipe (of clay)’ (cf. Hadiyya *gaawwa* ‘water pipe’ and Amharic *gayya* ‘pipe’).

Evidence for initial PHEC ***y** is scarce. In addition to the above, there are PHEC [354] **yaaʔ-* ‘to flow, assemble’ and [274] **yeyy-* ‘wolf, jackal’, which is found in Burji *yeyyee* and Gedeo *yaayye*, while Sidaama has inserted initial **h-** in *hiyy-eessa*.

Similar to the PHEC glide ***w**, single non-initial ***y** is vulnerable in HEC. There is no evidence of a single intervocalic ***y** within the root in HEC, in accordance with the rule of intervocalic glides loss ((iii)b) ***y** → **∅** / **V_V**.

Such loss of single intervocalic **y** has been reported, for example, in Burji (Tesfaye Baye, 2015, p. 43) and Sidaama (Yri & Pepper, 2019). In Gedeo, there are cases of elision of stem-final **y** and **w** (Eyob Kelemework, 2015, p. 87).

The glide gemination of originally non-geminates (e.g., plurative suffix ***-uw-a** → ***-uwwa**) is likely a loss-blocking strategy in these languages (for Hadiyya, see Tadesse Sibamo, 2015, p. 23). Thus, besides the gemination of intervocalic ***w** discussed above, a similar one is tentatively posited ***y** → **yy** / **V_V**.

⁴⁸ Sasse (1982, pp. 102, 108) reconstructs it as a PEC stative prefix.

⁴⁹ The following derived forms could be proposed for Hadiyya, Kambaata, and Sidaama: **ay-y-e* (acc), **ay-y-i* (nom) ‘who?’.

The following example shows no **y** reflex in Kambaata and Gedeo because they used a different support verb *ʔ (§4.2.4.2), and Gedeo **d'** is part of the reflexive denominative (§4.2.5.5).

PHEC [31] *oo y-, *oo ʔ 'to cry' → Burji oy- 'cry, weep'; PNHEC *oo y-, *oo ʔ 'to cry' → Had ooy- 'to weep, mourn', Kam oʔ- 'cry, weep, mourn', Sid ooy-, ʔoʔʔo yi; oʔo y- 'cry', Ged ood'- 'wail, weep, mourn' (cf. Oro oo(y)- 'cry').

As Sasse (1982, p. 156) notes, this stem is a compound from PEC *ʔoo y- 'say oo'.

Epenthetic **-y-** across morpheme boundaries is well attested in HEC. It is also well-attested in other EC languages, at least in EOT. A discussion and more details can be found under §3.4.1.2.

The following three forms show heterogeneous correspondences in which Burji **VyC** corresponds to **VVC** elsewhere in HEC:

Burji *laymi*: PHEC [634] *leem- 'bamboo' → PNHEC *leem-, *leem-n- 'bamboo' → Had and Kam *leema*, Sid, Ged *leemma* (cf. Oro *leemana* 'reed, bamboo').

Burji *dureyfi*, *dur-ee-f*: PHEC [428] *dur- 'rich' → Kam *duuballaaffa*, Sid, Ged *dur-eessa*.

Burji *ogeyfi*: PHEC [236] *og-, *og ʔ-y-f- 'traditional doctor' → PNHEC *og-, *og-aʔ-y-f- 'traditional doctor' → Had *ogeeffa*, Sid *ʔogeessa*, *ʔogeette*, Ged *ogeeyyicca* 'wise man', *ogeeyitte* 'wise woman', *ogeeyy* 'wise men' (cf. Oro *ogeessa* 'wise, expert').

Example [428] is likely Oromo-related (see Oromo *dureessa* 'rich', with the suffix **-ssa**; Lamberti & Sottile, 1997, p. 348),⁵⁰ [236] already contains derivational **-y-**, and only [634] is a lexical case. Since no trigger for root-internal epenthetic **-y-** is available in [634], I treat Burji *laymi* as an isolated

⁵⁰ Compare also Had *lik'-aayyis-* 'lend': PHEC [392] *lik'- 'to lend' → Burji *lik'-ees-*; PNHEC *lik'- 'to lend' → Kam *leʔ-ees-* 'lend of what cannot be returned; lend (money)', Sid *lik'-iiss-* 'lend', Ged *lik'-eess-* (cf. Oro *lik'-eessa* 'lend (money)'). This shows that comparable glide-vowel sequences may also arise secondarily and are therefore not by themselves probative for reconstruction.

irregular development rather than as evidence for reconstructing proto ***VyC**.sequence (see §2.2.1.2 (PHEC ***ee**)).

The glides **w** and **y** are found in number-marking suffixes and occur epenthetically in various HEC languages in various HEC languages (see §4.1.1.2 and §3.4.1.2).

Summary

The limited evidence for non-initial ***w** and the systematic loss of intervocalic glides suggest these sounds were already unstable in PHEC. The contrast between relatively frequent **y** gemination and rare **w** gemination may reflect phonotactic preferences that differentiated the two glides beyond their basic articulatory properties.

2.1.5 Velars, Glottals, and the Pharyngeal

The velars, glottals, and pharyngeal consonants of PHEC show a distinct pattern of development across the HEC languages. This section examines the six consonants in this class: the velars ***g**, ***k**, and ***k'**, the glottals ***h** and ***ʔ**, and the pharyngeal ***ħ**. These sounds display both stable retentions and innovative developments in the descendant languages. The reconstruction of these consonants reveals important insights into the phonological structure of PHEC and the subsequent sound changes that have affected the individual HEC languages. Particular attention is given to the patterning of these consonants in different phonological environments and their participation in morpheme boundary alternation processes (§3.2, §3.3, §3.3.4).

2.1.5.1 Velars

Highland East Cushitic has three velars: the voiced **g**, the voiceless **k**, and the ejective **k'**.

PHEC ***g**

The regular **g** reflexes across all HEC languages in both initial and non-initial position support the secure reconstruction of PHEC ***g**. Burji, however, exhibits limited cases of **g** palatalization to **j** before **i** (§3.2.1).

90 A Lexical and Phonological Reconstruction of Highland East Cushitic

PHEC	[36] *gam-	'to bite'	[702] *hig-	'to return, pass'	[80] *gal	'to enter'
Burji	gam-	'to bite'	hig-	'exchange'	gal-	'enter'
PNHEC	*gam-ʔ- → *gaʔm-	'to bite'	*hig-	'to return, pass'	*gal-	'to enter'
Had	gaʔm-	'to bite'	hig-	'pass'	gar-	'pass the night'
Kam	gaʔm-	'to bite'	hig-	'pass'	gal-	'pass the night'
Sid	gaʔm-	'to bite'	hig-	'return'	gal-	'enter, pass the night'
Ged	gaʔm-	'to bite'	hig-	'return'	gal-	'pass the night'

Additional evidence for PHEC ***g** includes: PHEC **[350]** **sang-* 'ox, bull',⁵¹ PHEC **[126]** **gog-* 'skin, hide', PHEC **[67]** **ang-* → PNHEC **ang-* 'hand, arm', PHEC **[707]** **magan-* 'sky, God', PHEC **[223]** **eg-* 'to wait', PHEC **[85]** **arg-* 'to send', and PHEC **[438]** **zuug-* 'to scrape, tan (hide)'.

Burji has limited instances of palatalized **j** corresponding to **g** in the rest of HEC before the high front vowel **i**. Some of them have a palatalized parallel in Oromo, where the palatalization is regular.

[400] Burji *jiira* 'fire': Kam *giira-ta* 'fire', Sid, Had, Ged *giira* 'fire'.⁵²

[701] Burji *jisk-*, *jig-* 'to demolish/destroy': Had *giggiss-* 'to shake off', Kam *giggiff-* 'shake off, let fall down', Sid *giggiff-* 'tear down part of a house in order to renew it, snatch, kick off', Ged *giggis-* 'stutter'.

[183] Burji *jibb-* 'to hate, neglect': Had *gib-* 'flee', Kam, Sid, Ged *gib-* 'refuse', Oro *jibba* 'dislike'.

It is likely a contact-induced change applied irregularly in Burji, because **[400]**, **[701]**, and possibly **[183]** are the only three established examples that support the positing of a rule. Oromo **[183]** *jibba* 'dislike' with **j** is not the only such parallel. Here, there is a chance that it is borrowed into Burji. Clearer loans with **j** into Burji and other HEC languages include:

[640] Oro *jabaa* 'strong', *jabeenna* 'strength', *jabeessa* 'make strong': Burji *jebeenna*, *jebuntaa* 'strength': Sid *jawaante* 'strength',

⁵¹ As Sasse (1982, p. 163) notes, **[350]** is a widespread word, but its origin is difficult to determine (see Amharic *sānga* and Oromo *sang-aa* 'ox for beef').

⁵² Probably related to **[940]** Burji *jirmalle* 'acacia' and Had *giraar* 'acacia tree, mimosa'.

Ged *jabeessa* 'strengthen' (cf. Arbore *goob-* 'become a strong young adult', Amharic *jäbdu* 'act of bravery').

[265] Oro *hujj-ii* 'work': Burji *hujee* 'work': Kam *huji-ta* 'work', Ged *huja* 'work'.

There are isolated instances of palatalized *j* in non-initial position [18], sometimes alongside the non-palatalized variant [708], discussed in §3.2.1. However, the regularity of *g* palatalization in Burji is further challenged by the fact that the only example that goes as far as PEC [4] shows non-palatalized *g* before *i* in Burji, although Oromo (and EOT languages) exhibit palatalized *j*.

[4] Burji *gilba* 'knee': Had *gurubbo* 'knee', Kam *gulubi-ta* 'knee', Sid *gulube* 'knee', Ged *gulubo* 'knee', Oro *jilba* 'knee'.

[632] Burji *giddee* 'inside': Had *guda* 'middle, half', Sid *giddo* 'inside', Ged *giddo* 'inside, in', Oro *jidd-uu* 'middle', *goda* 'inside, *gadi* 'interior'.

These examples point to an irregular Burji-specific palatalization of **g* → *j* before high front vowels, particularly *i*. Since the NHEC cognates consistently retain *g*, this change must post-date the split of Burji from PNHEC. Oromo is the most plausible contact source, because comparable palatalization is regular there, and Burji also shows clear Oromo loans with *j* in this section. The wider EOT parallels also provide a typological support for this established development in LEC.

While the voiced velar **g* shows contact-induced instability in Burji, the voiceless velar **k* exhibits a different pattern of distribution, particularly in non-initial position.

PHEC **k*

PHEC **k* is continued regularly in initial position. However, it mostly occurs as a geminate or in a cluster in post-vocalic position. Here are examples of the securely reconstructed PHEC **k* reflexes in initial position:

92 A Lexical and Phonological Reconstruction of Highland East Cushitic

PHEC	[705] *kul-	'to tell'	[418] *ku, *ko	'this' (m.nom, acc)	[30] *kaʔ-	'to get up, arise'
Burji	kulʔud	'to stutter'	ku, ka	'this' (m.nom, acc)	kaʔ(?)	'to get up'
PNHEC	*kul-	'to tell'	*ku, *ko	'this' (m.nom, acc)	*kaʔ-, *keʔ-	'to get up, arise'
Had	kur-	'tell'	kuk, ka	'this' (m.nom, acc)	kiʔ-	'to get up'
Kam	kul-	'tell'	ku, ka	'this' (m.nom, acc)	keʔ-	'to get up'
Sid	kul-	'tell'	kunni-, kónne	'this' (m.nom, acc)	kaʔ-	'to get up'
Ged	kul-	'tell'	kunni, konne	'this' (m.nom, acc)	kaʔa, keʔa	'to get up'

Two forms **[30]** **kaʔ-* and **keʔ-* 'to get up, arise' are reconstructed for PNHEC because, in addition to vowel ablaut, the following glottal stop also causes variation like that seen in the three different vowel reflexes **a**, **i**, and **e** (see §2.2.2.1).⁵³

Additional examples of initial **k** include: PHEC **[652]** **kiil-* 'to do magic', **[478]** **kott-* 'hoof, foot', **[327]** **kor-* 'saddle', **[391]** **kum-* 'thousand', PNHEC **[670]** **kaas-* 'to plant', and **[602]** **kin-* 'stone, rock'.

There is no evidence for a fully retained single non-initial ***k** (see Kambaata (Treis, 2008, p. 26)). It survives only in clusters (mostly **nk**, **rk**) and as a geminate (**kk**) in non-initial position.

PHEC **[51]** **lokk-*, **lekk-* 'foot, leg' → Burji *lukka*; PNHEC **lokk-*, **lekk-* 'foot, leg' → Had *lokko*, Kam *lokka-ta*, Sid *lekka*, Ged *lekka* (see Oro *luka* 'foot, leg').

PHEC **[161]** **bark-* 'head-support, pillow (of wood)' → Burji *bórk-e*; PNHEC **bark-* 'head-support, pillow (of wood)' → Had *barkuma*, Kam *borkaanu*, Sid *barko*, Ged *barko* (cf. Oro *bark-oo* 'headrest').

PHEC **[69]** **midk-* 'right(-side)' → Burji *mirga* 'right(-side)'; PNHEC **midk-* 'right(-side)' → Had *makka*, Kam *makki-ta*, Ged *midda*.

PNHEC **[397]** **kikkil-* 'armpit' → Had *kikkiro*, *kilkilo*, Kam *kikkili-ta*, *kilkilli-ta*; Sid *kikkiʔla*, *kikkilicco*.

⁵³ The origin of the glottal stop is a pharyngeal, which causes vowels to lower in some Cushitic languages. However, this is not the case in HEC, since the PHEC reflex of the pharyngeal ***ɕ** is the glottal stop ***ʔ**.

In [69], the **d** has assimilated to the following **k** in Hadiyya and Kambaata **madk-* with vowel assimilation/ablaut and to the preceding **d** in Gedeo via **mid-g-*. Oromo *mirga* ‘right(-side)’ likely influenced the Burji word with **r** and **g**. In the other examples [397], the root is reduplicated partially or fully in each language, resulting in single **k** and geminate **kk**.

There are morphophonological processes that contributed to the lenition and subsequent loss of potentially retained single non-initial **k** discussed in §3.2.2 and occasional palatalization in Burji §3.2.1.

PHEC *k’

PHEC *k’ is only arguably reconstructable, and its historical interpretation remains unresolved. Below is a summary of the attested correspondence series covered in the analysis and their frequency:

	Burji	Had	Kam	Sid	Ged	Frequency
1.	k’	k’	k’	k’	k’	Common
2.	k’	∅	∅	∅	∅	Less Common
3.	k’/k	k	k/g	g	k’	Uncommon

Note that these series are chiefly attested in initial position, and the reflexes are mostly sporadic in the second and third series. At present, the evidence does not support a single regular account of all three series. Series 1 favors reconstructing PHEC *k’, whereas series 2 and 3 may reflect differential loss, analogical reshaping, or an earlier conditioning that is no longer recoverable from the surviving forms. As with the other non-pulmonics discussed in §2.1.2, the discussion below evaluates both the HEC distribution and the wider EC context.

In NHEC, inherited *k’ is much less stable than Burji alone would suggest. Many NHEC words with **k’** lack corresponding non-pulmonic cognates elsewhere in EC and are therefore difficult to project back beyond HEC with confidence. Some may reflect reestablished PNHEC *k’, including early proto-level borrowings or post-split innovations (see §2.1.1.2 (PHEC *j)); others may continue older PHEC *k’ that was later lost outside NHEC or elsewhere in EC. For that reason, the relation between PHEC *k’ and PNHEC *k’ is better treated as unresolved than reduced to a single history.

	[189]		[392]		[108]	
PHEC	*k'araar-	'be bitter'	*lik'-	'to lend'	*sunk'-	'to kiss'
Burji	k'araar-	'be bitter'	lik'-ees-	'to lend'	sunk'-	'to kiss'
PNHEC	*k'araar-	'be bitter'	*lik'-	'to lend'	*sunk'	'to kiss'
Had	k'araar-	'be bitter'	lik'-aayyis-	'to lend'	sunk'-	'to kiss'
Kam	k'araar-	'be bitter'	leʔ-ees-	'to lend'	sunk'-	'to kiss'
Sid	k'araar-	'be bitter'	lik'-iiss-	'to lend'	sunk'-	'to kiss'
Ged	k'araar-	'be bitter'	lik'-eess-	'to lend'	sunk'-	'to kiss'

Two sets stand out because they preserve initial **k'** in NHEC together with broader EC support. One is [189], which occurs with **k'** in Dahalo *k'araare* 'bitter' and **q** in Somali *qaraar* and Maay *qereer* 'bitter'. The other is [115], where Gedeo, Hadiyya, and Sidaama *k'iid-* 'become cold' has multiple EC parallels with initial **k'** or **q** as well. In [392], by contrast, Kambaata shows a puzzling **ʔ** reflex instead of **k'**. This is not an isolated anomaly: **k'** is elsewhere replaced by **ʔ** or **ø** in Kambaata and the other NHEC languages. At least some of these losses may therefore reflect an earlier conditioning that is no longer recoverable from the surviving lexicon.

For more examples of PHEC ***k'** see PHEC [675] **k'oom-* 'calabash', [449] **k'al-* 'give birth, beget (of animals)', [582] **hak'k-* 'wood', [240] **k'ort'umm-* 'fish' (← Oromo), PHEC [116] **k'ok'k-* 'throat' → PNHEC **k'oonk'-* 'throat', and PNHEC [494] **k'ob-*, **k'ok'~k'obb-* 'to surround, to bend', [49] **k'uub-* 'egg', PNHEC [21] **c'ak'as-* 'to listen, spy'.

The non-pulmonic velar ***k'** discussed here contrasts with glottalized velars arising through fusion. In PNHEC, sequences of velar obstruent plus glottal stop undergo fusion to yield geminate **k'k'** (§3.3.1.3). These glottalized consonants appear in derived forms where the glottal stop belongs to a suffix or stem-final or additional suffix segment, whereas non-pulmonic ***k'** occurs in stem-initial and non-initial positions independently of morphological structure.

The examples below show initial **k'** in Burji corresponding to **ø** in NHEC. In contrast to the above examples, most of the following have parallels in the rest of EC, where **k'** and other correspondences are also attested. Therefore, at least some of them are arguable retentions in Burji.

[29]	k'	∅	Gloss
PHEC	*k'ab-		'to hold/seize'
Burji	k'af-		'to have'
PNHEC		*ab-	'to hold/seize'
Had		amad-	'hold, seize'
Kam		ap'p'-	'catch fire, hold, seize'
Sid		amad-	'hold, seize'
Ged		abid-	'hold, seize'
cf. Oro	k'aba		'have, possess, seize'

In [29], Burji and Oromo **k'** corresponds to **k'** in numerous EC languages:

Dir *k'ap*- 'catch, have', Mos *k'aba* 'catch', Boni Kilii *k'aba?*- 'to catch, seize', Arb *k'abad*'- 'seize', and Gaw *gap*- 'catch, seize'.

Saho *ab-it* 'take for oneself', Afar *-ibbid*- 'catch'; Bay *abat*- 'catch, seize personally', Boni Bireeri *owd*'-, Boni Jara *obed*'- 'to catch, seize', Elm *ap* 'to have, to seize'.

The initial **∅** in PNHEC **ab*- 'to hold/seize' also has **∅** correspondences in Saho, Afar, Bayso, Boni, and Elmolo.

[74]	k'	k, g	∅	Gloss
PHEC	*k'ub-	*-kub-		'to fall (down)'
Burji	(k'up'an-)	(gambun?-)		'to squat, to fall by itself on the face'
PNHEC			*ub-	'to fall (down)'
Had			ub-	'fall (down)'
Kam			ub-	'fall (down)'
Sid			ub-	'fall (down)'
Ged			ib-	'fall (down)'
cf. Oro		kufa		'fall (down)'

In [74], the two reconstructions **k'ub*- and **-kub*- 'to fall (down)' remain competing possibilities. If one accepts the EC-supported reconstruction **-kub*- 'to fall (down)', PNHEC **ub*- would require loss of the initial velar under a conditioning that is no longer directly observable.⁵⁴ If one instead privileges

⁵⁴ The EC parallels are Rendille, Somali, Boni *kuf* 'to fall', and Yaaku *kup*- 'die'. One assumption may be a remnant prefix verb root preceded by a vowel that triggered the lenition and elision of the **k** (§3.2.2).

the Burji form, the set joins the broader **k'** ~ **∅** series but remains irregular. The set, therefore, does not resolve the history of ***k'**; rather, it shows that the correspondence cannot yet be reduced to a single regular development.

The correspondence of non-pulmonic **k'** with the pulmonic **k** is not an isolated case, as discussed below.

[71]	k'	∅	Gloss
PHEC	*k'utañ-		'thorn'
Burji	k'uwwa, k'uwa		'thorn'
PNHEC		*utañ-	'thorn'
Had		utta	'thorn'
Kam		uta	'thorn'
Sid		uta	'thorn'
Ged		ute	'thorn'
cf. Oro	k'orraattii		'thorn'

The EC parallels of **[71]** with initial **k'/q** include:

Dir *k'udd-éét* 'thorn, words which prick', Mos *k'uddēca*, Som *qodañ*, Maay *qurun*, Arb *k'ed'eh*, Ts'am *q'ants'-e* 'thorn'.

Various other correspondences are found in Konso *faatta/fa?atta*, Saho *keena* 'thorn, prickly', Afar *keena*, Rendille *kuḏáñ*, and Boni *upa(h)* 'thorn'.

[381]	k'	∅	Gloss
PHEC	*k'ul-, *k'ul~k'ul~l-?-		'leech'
Burji	k'ulaallee		'leech'
PNHEC		*ul-, *ul~ul~l-?	'leech'
Had		urulla	'leech'
Kam		urulla	'leech' ⁵⁵
Sid		ulaa?ulle	'leech'
Ged		ula?ulla	'leech'
cf. Oro		ulaan?ul-a	'leech'

In addition to Oromo *ulaan?ul-a* 'leech', the only available EC parallels of **[381]** are Afar *ṣallullé*, Bayso *halalac'e*, and Somali *ṣalaaṣul* 'leech', with none

⁵⁵ The Kambaata **r** reflex of ***l** is probably due to Hadiyya influence, where such a change is regular.

supporting the reconstruction of PHEC **k'**. Even so, the PHEC reconstruction is retained only provisionally. Burji **k'** places these sets within the broader initial **k'** ~ **∅** pattern, but the wider EC evidence is too uneven to decide whether the Burji forms are genuine retentions, secondary preservations under contact influence, or mixed cases. The corresponding Oromo **k'** in two of the four sets may reflect contact influence, but it does not by itself determine directionality. What is clearer is that PNHEC shows initial **∅**. Accordingly, the PNHEC reconstructions are kept with **∅**, while the PHEC reconstructions with ***k'** are treated as provisional and unresolved. Some may reflect genuine PHEC retentions later lost elsewhere in HEC or EC; others may not.

There is a marginal series of **k'**, **k**, and **g** correspondences (see [74] above) attested in the following few examples:

[47]	k'	k	Gloss
PHEC	*k'uf-		'to cough'
Burji	k'ufey-		'to cough'
PNHEC		*kuf-	'to cough'
Had		kut't'-	'cough'
Kam		kuc'c'-	'cough'
Ged	k'ufa-aʔ-		'cough'
cf. Oro	k'ufaʔa		'cough'

Here are the EC parallels in set [47]:

Dir *k'ayy-ood-*, Kon *k'aaw-ad-*, Som *qufaʕ*, Maay *quffi* 'cough', Boni *k'ufaʔ-*, Das *ɣuf-ume*, Yaaku *qopeʔe-*, Ts'am *ɣufaʕ* 'to cough'.

Saho *ufuʕe* 'to cough', Bay *ufii-* 'cough', Boni *'ufaʔ-*, Elm *úfe*, Arb *ʔuf(e)ʔ-*, Das *ʔúfuk* 'to cough'.

PHEC **k'** corresponds to **k'/q/g** in most of the EC parallels and, as usual, **∅** in Saho, Bayso, Boni, and Elmolo in addition to Arbore and Dasanach. Rather than **∅**, PNHEC shares **k** with Afar *kaʕite*, *kaʕùw-eñʕe* 'cough'.

98 A Lexical and Phonological Reconstruction of Highland East Cushitic

[115]	k'	k	g	Gloss
PHEC	*k'ab y-t-			'become cold'
Burji		kabb-		'cool, become cold'
PNHEC	*k'iid-			'become cold'
Had	k'iid-			'cool, become cold'
Kam			gid-	'cool, become cold'
Sid	k'iid-			'cool, become cold'
Ged	k'iid-			'cool, become cold'
cf. Oro	k'abb-an-			'become cool (liquid)'

As in the previous set [115], PHEC ***k'** corresponds to EC **k'/q**:

Dir *k'appann-aw-* 'be cool (weather or personality (calm))', Mos *k'appannayya, k'apana* '(be) cold', Kon *k'ap-p-ann-aaw-* 'become cool (liquid)', Ren *xóbo* 'coolness, cool', Som *qabow* 'coldness, cold', Maay *qoboob* 'cold', Arb *k'abbat* 'cold (thing)'.

Bay *ambalki* 'cold, low temperature', Boni *ab'óob* 'coolness, cool'; Das *ʔabbúnab* 'cold', Elm *ʔápat* 'shadow; become cold'.

Even though PNHEC has ***k'**, Bayso, Boni, and Elmolo have **ø** as usual.

[288]	k'	g	Gloss
PHEC	*k'el-		'snail'
Burji	k'ilaank'illoo	k'ilangiloo	'snail'
PNHEC	*k'el-		'snail'
Had	k'elema, k'eloʔo		'snail'
Kam	k'eloma-ta		'snail'
Sid	k'erek'enicco	gerk'eʔnâma	'snail'
Ged	k'ak'ank'ale		'snail'

There are no solid EC parallels for [288] (see Oromo *c'iilalluu* 'snail').

The irregular reflex **g** is attested in Burji [288] *k'ilangiloo* 'snail', Sidaama *gerk'eʔnâma* 'snail', and Kambaata [115] *gid-* 'cool, become cold'. Tesfaye Baye (2015, p. 79) notes that in Burji, "/k/ is sometimes found in free variation with /g/ based on dialectical differences. People from the northern area of Burji and the surrounding of Soyama use the variant /k/ while those from the southern use /g/...". This explains the above **k' ~ g** variation in [288]. Such variation has also been reported by Sasse (1982, p. 18).

The velar consonants ***g** and ***k** show relatively stable development across the HEC languages. By contrast, the non-pulmonic ***k'** remains an unresolved problem with a series supporting an arguable reconstruction of PHEC ***k'**, but two series that do not yet yield a single regular explanation. Some mixed reflexes may continue an older conditioning that is no longer recoverable; others may reflect differential loss, analogical reshaping, or contact. Likewise, PNHEC ***k'** cannot be assigned a single origin: some forms are likely reestablished within PNHEC, while others may preserve older PHEC lexicon lost elsewhere in EC. For the present, the most defensible conclusion is therefore that PHEC ***k'** is arguable but not resolved, and that the three-way contrast among PHEC ***g**, ***k**, and ***k'** is best treated as a provisional working reconstruction rather than a fully regularized historical result. The distributional patterns indicate that ***k** frequently occurs as a geminate or in consonant clusters in non-initial positions, while ***k'** shows a limited distribution in non-initial positions.⁵⁶ These observations align with broader tendencies in East Cushitic historical phonology, particularly the retention of initial consonant contrasts across the language family.

2.1.5.2 Glottals

The two HEC glottal consonants are **h** and **ʔ**. Elision or lack of non-initial glottal consonants (intervocalic, syllable coda, in a cluster, etc.) has been reported by Sasse (1982, pp. 18, 35) and Tesfaye Baye (2015, p. 61) for Burji, by Eyob Kelemework (2015, pp. 46, 83–86) for Gedeo, by Tadesse Sibamo (2015, p. 34) for Hadiyya, and by Anbessa Teferra (2000, p. 27) for Sidaama. Similar to the sonorants, the glottal consonants **h** and **ʔ** are also used for epenthesis in HEC (§3.4.2).

PHEC ***h**

Initial PHEC ***h** is continued in all HEC languages, so its reconstruction is secure.

⁵⁶ Irregular non-initial **k'(k')** should be evaluated against three possible sources: direct inheritance, contact-induced variation, and fusion as a result of boundary alternations. The fusion path is particularly relevant for forms showing geminate **k'k'** in PNHEC post-vocalically, as these systematically arise from velar-glottal sequences (§3.3.1.3) rather than continuation of inherited geminate **k'k'**.

100 A Lexical and Phonological Reconstruction of Highland East Cushitic

PHEC	[228] *harr-	'donkey'	[702] *hig-	'to return, pass'	[380] *hik'k'-, *hek'k'-	'have hiccups'
Burji	harree	'donkey'	hig-	'exchange'	hirc'imm eed'-	'have hiccups'
PNHE	*harr-	'donkey'	*hig-	'to return, pass'	*hik'k'-, *hek'k'-	'have hiccups'
C					hik'k'a?-	'have hiccups'
Had	halla	'donkey'	hig-	'pass'	hik'k' a?-	'have hiccups'
Kam	harri-ta	'donkey'	hig-	'pass'	hek'k' i	'have hiccups'
Sid	harre	'donkey'	hig-	'return'	yaa hek'k'e? mef-at-	'have hiccups'
Ged	harre	'donkey'	hig-	'return'		

Set [380] also suggests Oromo influence. Its Oromo parallel, *hirk'inf-* 'have hiccups', contains **f**, whereas the corresponding HEC forms are irregular in this position: Burji has **r** (see §3.4.4) and Gedeo has **f**. Since Oromo **f** regularly reflects earlier ***s** → **f** (see §2.1.1.2 (PHEC ***f**)), probably in a causative/derivational suffix in **-s**, the Gedeo **f** is best treated here as contact-aligned rather than as an inherited regular reflex.

Further examples of PHEC ***h** include: PHEC [492] **haar-* 'new', [256] **ham ?* 'to slander', [768] **hab-* 'to forget', [461] **has-* 'to want, seek', [493] **hii?-* 'take!', PHEC [722] **hoog-* 'to clean', PNHEC [700] **ha?m-i-jf-* 'ensete tree root/corm', and [601] **hos-* 'to pass/spend the day'.

Overall, reflexes of reconstructable **h** are mainly attested before the low vowel **a**, less commonly before the front vowels **e** and **i**, rarely before **o**, and lacking before the back vowel **u**.⁵⁷ The less common occurrence of **h** before vowels other than **a** might have contributed to an irregular analogical dropping of initial **h** before front vowels in some HEC languages, e.g., in Kambaata *iinc'a?-* 'stretch (body)' ← PHEC [390] **hiit'-* 'to stretch', *ĩ* 'take!, here' ← PHEC [493] **hii?-* 'take!'.

The HEC languages have largely lost non-initial ***h**. The relevant morphophonological details are in §3.2.2, where the lenited **h** from PHEC ***k** is discussed.

⁵⁷ There are instances of Burji **h** corresponding to NHEC **w**, which could be taken as evidence for reconstructing ***hu**. However, there is also room to argue for a merger with ***hu** and a reconstruction of ***h** for such series.

PHEC *ʔ

In HEC, the glottal stop ʔ appears as a retention of *ʔ in monosyllabic stems and serves multiple morphological functions. It serves as a reflexive and genitive suffix, and is also found as an epenthetic sound. Below are reflexes of the securely reconstructable non-initial *ʔ, which are presumably part of the stem:

	[431]		[34]		[293]	
PHEC	*raʔ-	'ripen, be ripe'	*baʔ-	'be lost'	*waʔ-	'water'
Burji	raʔ-	'ripen, be ripe'	bey-	'be lost'	waa	'water'
PNHEC	*raʔ-	'ripen, be ripe'	*baʔ-	'be lost'	*waʔ-	'water'
Had	liʔ-	'ripen, be ripe'	biʔ-	'be lost'	woʔo	'water'
Kam	reʔ-	'ripen, be ripe'	baʔ-	'be lost'	woʔa	'water'
Sid	leʔ-, raʔ-	'ripen, be ripe'	baʔ-	'be lost'	wa, waa	'water'
Ged	raʔ-	'ripen, be ripe'	baʔ-	'be lost'	wedeʔe	'water'

As discussed in §2.2.2.1, the vowel variation before a glottal stop is found even within the same language in the above examples: Sidaama **[431]** *leʔ-*, *raʔ-* 'ripen, be ripe', and Burji **[34]** *bey-* (Roba Dame & Wedekind, 2008, p. 8), *be-*, *bay-* (Hudson, 1989, p. 180) 'be lost', and Gedeo **[293]** *wedeʔe* (Wedekind, 2008, p. 147), *wadaʔe* (Hudson, 1989, p. 165) 'water'.⁵⁸ The **d** and **e** in Gedeo **[293]** *wedeʔe* 'water' might be suffixes, while ʔ is epenthetic (§3.3.4.1). Although vowel variation following **w** is common, at least in NHEC, the glottal stop is the actual trigger of this vowel alternation, as demonstrated by the fact that this phenomenon includes Burji and all three examples (§2.2.2.1).

Post-vocalic ʔ in Burji has different correspondences conditioned by boundary alternation processes (§3.3.3). Burji **[431]** *raʔ-* 'ripen, be ripe' has a preconsonantal stem *raa-* (Sasse, 1982, p. 157) since ʔ is dropped before consonants with concomitant lengthening of the preceding vowel (Sasse, 1982, p. 18).⁵⁹ However, the causative stem of **[34]** Burji *bey-* 'be lost' is *ba-ss* (caus) (Sasse, 1982, p. 34) without vowel doubling, similar to *t'ay-* 'to extinguish' becoming *t'as-s-* in the causative but with the fusion of ʔ and **s** into **ss**. This means that vowel doubling is not regularly applied in Burji when ʔ is dropped.

⁵⁸ Burji also has *wáyu* (*wéu*) (subj), *way-i-n-* (gen) 'water' (Sasse, 1982, p. 186).

⁵⁹ Another example is Burji *raʔeed'aa(ga)* (adj.m), *raaʔettaada* (adj.f) 'cooked, ripe'.

Intervocalic ʔ is scarcely found in Burji (Tesfaye Baye, 2015, p. 30).

PHEC [144] **meʔ*- ‘how many? how much?’ → Burji *miʔa* ‘how many? how much?’.

PHEC [493] **hiiʔ*- ‘take!’ → Burji *heʔee* ‘take!, here’, *heettaa* (interj) ‘Here it is!’.

Otherwise, non-initial ʔ in Burji is found as a suffix (§§4.2.2.1, 4.1.2) and following sonorant consonants, in contrast to the NHEC languages (§3.3.2). There are three regular morphophonological processes concerning ʔ in Burji, discussed in detail in the respective sections:

- (a) elision with compensatory vowel doubling (§3.3.3),
- (b) **-y-** epenthesis (§3.4.1.2),
- (c) complete progressive assimilation (§3.1.1.3)

Therefore, in addition to fully retained ʔ, a glottal stop is reconstructed in non-initial position where ʔ is found in NHEC corresponding to compensatory vowel doubling, epenthetic **-y-**, or geminated suffix consonant in Burji. The evidence is almost exclusively found in monosyllabic stems. The instances of HEC glottal stop found in disyllabic or multisyllabic stems are mostly epenthetic and affixes.

The previous paragraphs examined the reflexes of PHEC *ʔ in non-initial position and determined their role as suffixes only. Initial ʔ, on the other hand, lacks consensus as to its phonemic status in HEC. Therefore, the remaining part of this section on PHEC *ʔ synthesizes the synchronic evidence documented for initial ʔ in the individual languages and argues why initial *ʔ cannot be reconstructed.

Most HEC grammars posit initial ʔ for vowel-initial roots, including Burji (Tesfaye Baye, 2015, pp. 47, 30, 31), Sidaama (Yri & Pepper, 2019), Hadiyya (Tadesse Sibamo, 2015, p. 33), and Kambaata (Treis, 2008, p. 26). Eyob Kelemework (2015, p. 41) postulates vowel-initial roots for Gedeo, conforming to Hudson (2007) but disagreeing with Wedekind (1985).⁶⁰

Historically, more than one consonant (e.g., PEC *ʕ and *kʰ) has merged into ʔ in initial position in HEC. The pattern of merging and influence on the

⁶⁰ There are differences between orthography and phonemic transcription (see Yri & Pepper, 2019).

following vowel, if any, is unclear. Thus, vowel-initial roots are reconstructed in this study without a default glottal stop.

Further arguments in favor of vowel-initial stems include the existence of vowel-initial suffixes in Kambaata, Gedeo, and Hadiyya (e.g., case vowels). In Kambaata, there is no evidence of a glottal stop when vowel-initial nominals are reduplicated or in compounds (Treis, 2008, pp. 27, 178).

Stem(s)	Gloss	Reduplicate/ Compound	Gloss
ʔissoʔóot	'they'	ʔísissoʔóont	'they ... each other'
min- + ʔanná	'house' + 'father'	minanná	'householder'
abb-á + am-á-ta	'big' + 'mother'	abb-am-á-ta	'grandmother'
lank-í + am-á-ta	'second' + 'mother'	lankaam-á-ta	'sister of mother' ⁶¹

Contrary to the above analysis, Treis treats the initial glottal stop as a phoneme, as evident in the reduplicated initial syllable of *ʔontú* 'five' → *ʔóʔontú* 'five each' and she explains *ʔísissoʔóont* 'they ... each other', *ʔaʔn-aʔnoʔóot* (pl) 'you' (p. 358) as rather a reduplication of the first vowel and the subsequent consonant. In the compound *minanná* 'head of household', the explanation is that the stem-initial glottal stop of the second constituent is usually deleted. Those are three different analyses, so it might be convenient to instead posit that all of these stems begin with a vowel and assume epenthesis of ʔ in *ʔóʔontú* intervocalically (§3.3.4.1). Treis (2008, p. 358, fn. 34) acknowledges the lack of ʔ in the reduplicated *ʔísissoʔóont* 'they ... each other' and *minanná* 'head of household', saying they "might be considered an argument not to treat the glottal stop as a phoneme but as a default word-initial phone".

The vowel-initial stems are also evident in Gedeo's phrasal expressions.

Stem(s)	Phrase	Gloss
giddo ʔuggi	giddougg	'put it inside!'
ʔayyi -ʔa	ʔayyia	'whom-for?'
hoola ʔabbo	hoolabbo	'oh dear!'

The lack of an intervening consonant results in diphthongs (**ou** and **ia**) when the vowels are different, as in the first two examples, and elision of one vowel when they are identical, as in the third example (Wedekind, 1980, p. 161). Rather than taking the lack of the glottal stop in such forms as evidence for

⁶¹ Additional examples include: *abb-á* 'big' + *ann-á* 'father' → *abb-ann-á* 'grandfather' and *lank-í* 'second' + *ann-á* 'father' → *lankaann-á* 'brother of father'.

vowel-initial stems, Wedekind describes the above examples as instances of initial glottal stop elision.

In Hadiyya, as well, the vowel-initial stems mostly show no glottal stop in their compound formations.⁶²

Stem(s)	Gloss	Compound	Gloss
moólllo + ?ánna	'ethnic group' + 'father'	moollánna	'lord/god'
?allaácca + ?uúlla	'keeping' + 'land'	?allaccuúlla	'pasturage'
?áma + ?aáyya	'mother' + 'sister'	?amaáyya	'aunt'
?ánga + ?oóbbba	'hand' + 'palm'	?angoóbbba	'palm'
máto + ?ille	'one' + 'eye'	matille	'one-eyed person' ⁶³

There are three Hadiyya examples with a glottal stop in compounds:

Stem(s)	Gloss	Compound	Gloss
mine + ?áma	'house' + 'mother'	mi?náma	'householder (mother)'
mine + ?ánna	'house' + 'father'	mi?nánna	'householder (father)'
?eeliínco + ?aágga	'sun' + 'entrance'	?eeliinci?aágga	'morning'

These examples probably involve complex genitive constructions (see Genitive glottal stop ? in §4.1.3.1).

In sum, despite the deficient distribution, the “/?/ is a fully-fledged consonant with all the ‘human rights’ of any consonant” (Yri, 2004). Beyond its status as a retained proto-phoneme, PHEC *? plays a morphophonological role in PNHEC. The glottal stop triggers fusion when adjacent to obstruents, yielding geminate glottalized consonants (§3.3.1). This morphophonological function makes *? central to understanding the distribution of consonants in modern HEC languages, as inherited non-pulmonics should be distinguished from the glottalized geminates derived from fusion rather than from direct inheritance. Further discussion of the non-initial glottal stop in morpheme boundary alternations is detailed in §3.3 and in morphology in §4.1.4.1, §4.1.3, §4.2.3.1, §4.2.5.1 and §4.2.4.2.

⁶² It is unclear whether these forms are compounds, as Tadesse Sibamo (2015) writes them as a single form, whereas Ritter (2007) presents them in two separate forms.

⁶³ Additional examples include: *lob-* ‘great’ + *?ánga* ‘hand’ → *lobángo* ‘helpful, valuable’, *?ánna* ‘father’ + *?áre* ‘wife’ → *?annáre* ‘step-mother’, and *?ánna* ‘father’ + *?abbáyyo* ‘brother’ → *?annabbaájjo* ‘uncle’. Data source: Tadesse Sibamo (2015).

2.1.5.3 The Pharyngeal

PHEC *ħ

The correspondence series Burji **h/ø** and NHEC **h/w/ø**, conditioned by adjacent vowels and word position, requires explanation. Two analyses compete: (1) an analysis as epenthesis, treating these segments as phonologically inserted to resolve hiatus, or (2) the reconstruction of PHEC *ħ (a voiceless pharyngeal fricative) with regular reflexes and some sporadic cases. The reconstruction analysis gains support from EC correspondences and reflex distribution patterns. The following data establishes the correspondence patterns, followed by the details of each interpretation:

PHEC	[520] *ħuy ʔ-m	'be full'	[163] *ħur-ʔ-	'young ox, bull'	[519] *ħilili-	'smoke'
Burji	huum-	'become full'	(h)arʔeyi	'ox'	hilili	'smoke'
PNHEC	*ħuy ʔ-m-	'be full'	*ħur-ʔ-	'young ox, bull'	*ħilili-	'smoke'
Had	woʔm-	'be full'	woʔla	'calf'	wiriira, wuriira	'smoke'
Kam	wiim-, yiim-	'be full'	worjaamu	'young bull'	wilili-ta	'smoke'
Sid	woʔma, hom-	'all, full'	war-aamo	'ox, bull'	wiliile	'smoke'
Ged	woʔm-	'be full'	-	-	wiliille	'smoke'

The reflexes show a consistent pattern: Burji changes *ħ to **h** initially, while the NHEC languages develop **w** before high vowels. The reflex **y** is attested dialectally in Kambaata [\[520\]](#) alongside **w**. Treis (2008, p. 39) found that, in contrast with the Duuraame dialect, Daambooyya replaces **w-** with **y-** before long front vowels.

Duuraame	Daambooyya	
<i>wiimú</i>	<i>yiimú</i>	'to fill'
<i>wiitú</i>	<i>yiitú</i>	'to grind'
<i>weeffú</i>	<i>yeeffú</i>	'enset plant' (sg)

Usually, the **h** reflex is optionally dropped before the vowel **a** in Burji, as in [163] *(h)arʔeyi* ‘young ox, bull’.⁶⁴ Sidaama *hom-* ~ *wom* ‘be(come) full’ is further evidence for positing a unique phoneme for such a series.

The pattern observed in [520], [163], and [519] is confirmed by additional attestations in the same environment. Before ***u**, Burji consistently shows **h** and NHEC shows **w**: PHEC [607] **huy* ʔ-s- ‘to fill’, [207] **hun-*, **hut-* ‘strength, power’, [208] **hur ab-* ‘to pour’, and [1030] **huk-* ‘be(come) thin’. Six examples establish this correspondence as secure.

Outside of initial position before the high vowels ***u** and ***i**, the reflexes show environment-dependent variation. The correspondence sets below illustrate these context-sensitive patterns, which require analysis of each vocalic environment separately.

PHEC [467] **hons-* ‘nine’ → Burji *wonfa* ‘nine’; PNHEC **hons-* ‘nine’ → Had *honso* ‘nine’, Kam *honsu* ‘nine’, Sid *honse* ‘nine’.

PHEC [698] **hokol-* ‘to roll’ → Burji *kooliy-*, *kocikools-* ‘to roll’; PNHEC **hok~kol-* ‘to roll’ → Had *onkoroll-*, *(h)onkoroll-* ‘to roll’, Kam *onkoloʔl-* ‘roll’, Sid *konkolati* ‘to roll’ (see Oro *konkolaad’d’a* ‘roll, roll down hill’).

Before other word-initial vowels (e.g., **o**), the only reasonable regularity is in the Hadiyya reflexes : [467] *honso* ‘nine’ and *(h)onkoroll-* ‘to roll’, suggesting NHEC **h** before **o** for ***h**. Otherwise, Burji is the one with the **w** reflex in Burji *wonfa* ‘nine’, an anomaly expected in a numeral. The only other example before **o** has even more sporadic reflexes: PHEC [698] **hokol-* ‘to roll’ with assimilation to the flowing velar in Sidaama, **h** ~ **ø** variation in Hadiyya, loss in Kambaata, and apheresis (loss of initial syllable) in Burji.

The non-initial reflexes exhibit different series of correspondences in the following cognate sets:

PHEC [439] **reñ-* ‘death’ → Burji *reya*; *reeya*, *reeha*, *oya* ‘death’; PNHEC **reñ-* ‘death’ → Had *leho* ‘death’, Kam *rehu-ta* ‘death’, Sid *reo*, *rewo*, *reo* ‘death’, Ged *reyo*, *reo* ‘death’.

⁶⁴ See the feminine set: Burji *hil-icco*, *hil-aanno*, Had *weeʔ-icco*, *waʔla*, Kam *w(aʔ)eccu-ta*, and Sid *waʔr-icco* ‘calf, young’ (Hudson, 1989, p. 35).

PHEC [184] **reñ* ? ‘die’ → Burji *re-*, *rey-* ‘die, be humiliated’; PNHEC **reñ* ? ‘die’ → Had *leh-* ‘die’, Kam *reh-* ‘die, fall (of prices)’, Sid *re-*, *re?* ‘die’, Ged *rey-*, *re(?)* ‘die’.

PHEC [329] **reñ-f-* ‘corpse’ → Burji *reesi*, *reesi* ‘corpse’; PNHEC **reñ-f-* ‘corpse’ → Had *leeḑa*; *leeḑḑa* ‘corpse’, Kam *re(e)ḑa* ‘corpse’, Sid *reeḑḑa*, *reeḑa* ‘corpse’, Ged *reenḑa*, *renḑa* ‘corpse’, Oro *reeḑa* ‘corpse’.

PNHEC [627] **zoñ* ? ‘to wander’ → Had *do?* ‘roam, wander’, Kam *zoh-* ‘emigrate’, Sid *do-* ‘encircle, go round, wander’, Ged *zoor-* ‘roam, wander’.

PHEC [12] **mañ-* ‘what?’ → Burji *miyyaa*, *miya* ‘what?’; PNHEC **mañ* ‘what?’ → Had *ma*, *maha* ‘what?’, Kam *ma*, *maha*, *maḑa* ‘what?’, Sid *ma*, *maa*, *mai* ‘what?’, Ged *maa* ‘what?’.

PHEC [443] **uñw-* ‘to give’ → Burji *u(u)ww-* ‘give’; PNHEC **uñw-* ‘to give’ → Had *uww-* ‘give’, Sid *ḑui*, *aa*, *uww-* ‘give’, Ged *uww-* ‘give’.

The non-initial reflexes, as seen in [627], [329], [184], [439], and [12], show mostly Burji *ø* (or epenthetic *-y-*), Kambaata/Hadiyya *h*, and Sidaama/Gedeo *ø* (or epenthetic *-y-/-w-*). Such traces distinguish PHEC **ñ* from PHEC **h*, with the latter being completely lost in non-initial position. The alternation *y* ~ *ø* in non-initial reflexes is documented across multiple sources, though transcription practices vary. The following list [439] demonstrates that this variation may reflect linguistic reality in Burji, Sidaama, and Gedeo rather than inconsistent documentation:

Burji *reya* (Roba Dame & Wedekind, 2008, p. 69); *reeya*, *reeha* (Hudson, 1989, p. 47) ‘death’,

Sidaama *reo* (Yri & Pepper, 2019; Hudson, 1989, p. 47); *rewo* (Indrias Xa'miso et al., 2007, p. 489) ‘death’,

Gedeo *reyo* (Wedekind, 2008, p. 33); *reo* (Hudson, 1989, p. 47) ‘death’.

This cross-source consistency confirms the alternation as a synchronic phonological process in Gedeo (Eyob Kelemework, 2015, p. 36), suggesting incomplete loss of the pharyngeal reflex.

External comparative evidence strongly supports reconstruction of PHEC ***h**. Somali, within EC, shows **h** precisely where HEC languages have the variable reflexes **h/ø/y**. This correspondence pattern across EC indicates inheritance of ***h** from PEC, preserved in Somali, Afar, Saho, and Dullay, but variably reflected in HEC:

	PHEC	Burji	Sidaama	Somali	Gloss
[68]	* malaḥ -	malaa	malaa	malaḥ	'pus'
[38]	* liḥ -	lia, liya	lee, leye	liḥ	'six'
[20]	* sazeḥ -	fadia, fadiya	sase	saddeḥ, sadeḥ	'three'
[71]	* k'utaḥ -	k'uwa, k'uwwa	uta	qodaḥ	'thorn'
[22]	* ḥanʔur -	(h)anʔuroo	urco	ḥundur	'navel'

Additionally, PHEC ***h** as the reflex of PEC ***h** explains the persistence of identical final **a** across all the HEC in cognates like *mala* 'pus' from (PHEC (and PEC) [68] ***malaḥ**- 'pus' (§2.2.3.1, Terminal -a).⁶⁵ The systematic correlation between HEC reflex variation and Somali **h** undermines an epenthesis-only account, which would not predict cross-linguistic correspondence.

However, two EC parallels show correspondences other than **h** for the regular initial series (1): PHEC ***h** → Burji **h**, NHEC **w** before **u**:

PHEC [163] ***ḥur-ʔ**- 'young ox, bull': Saho *awur* 'bull, animal for mounting', Afar *awur* 'bull, ox', Som *awr* 'male camel', and Das *ʔauric* 'non-castrated camel', Ren *óor* 'bull (of camel or cattle)', Maay *oqr* 'male camel', Boni *óor* 'male elephant', Arb *ʔáar* 'bull', Elm *aar* 'male crocodile', Bay *aar* 'bull, ox'.

PHEC [208] ***ḥur ab**- 'to pour': Bay *woraab*- 'drink water, be watered'; Ren *warab*- 'draw water', *wor* 'waterhole, borehole', Som *waraab-i*- 'to water', *wár(-tii)* pool, pond', Maay *war* 'swamp', Arb *war(a)b*- 'draw and fetch water', Das *wár* 'river (the Omo)', Kon *oraap*- 'fetch, dip, pour (water only?)', Yaaku *irp-a*- 'draw water'.

⁶⁵ The reflexes of the PEC voiced pharyngeal ***ʕ** do not show unique behavior when compared to the reflexes of PHEC ***ʔ**, e.g., PEC [30] ***kaʕ**- 'to get up, arise' → PHEC ***kaʔ**-, ***keʔ**- 'to get up, arise', so they have likely merged. PHEC [519] ***ḥilill**- 'smoke' has parallels in Afar *ʕir*, *ʕer* 'smoke' and *ʕeəri* 'steam' with an unexpected voiced pharyngeal **ʕ** other than **h**, so if they are cognates, an intermediate ***h** or irregular Afar **ʕ** must be posited to explain the PHEC correspondence.

Many of the reflexes are comparable to those of NHEC with **w**, and others show mostly **ø** or a rounded vowel, proving the establishment of the **u** environment. Since they are only two examples, their implications with respect to the PEC reconstructions will need to be considered, for instance: (a) they might be reflexes of a consonant other than ***h**, e.g., a voiced glottal fricative ***ɦ**, or (b) they might be examples of ***h** change conditioned by the following **u** as occurs in NHEC. For now, the latter hypothesis is satisfactory for addressing the HEC reflexes, given the limited parallels with EC. However, a wider EC reconstruction might require considering the additional phoneme hypothesis if supported by further evidence. Here is one comparative set supporting the reconstruction of ***h** to address such a challenge:

PHEC [139] **wozan-* 'heart' → Burji *wodana* 'heart'; PNHEC **wozan-* 'heart' → Had *wodano*, *wadno* 'heart', Kam *wozana*; *wazana* 'heart, mind', Sid *wodana* 'mind (heart), human heart', Ged *onna* 'heart, courage', Oro *onn-ee* 'heart, courage', Kon *otan-ta* 'center', Saho *wazan-a* 'heart', Bay *wadan* 'heart', Ren *w'éyna* 'heart', Som *wadne*, *wadane* 'heart', Maay *wenni* 'heart', Boni *w'éene* 'heart', Das *^huođin-ni* 'heart'.

The reconstruction of PHEC **wozan-* 'heart' is without issue as long as the current hypothesis of Oromo influence on Gedeo *onna* 'heart, courage' is accepted as an explanation for the unexpected loss of initial **w**. Even in that case, however, the Oromo cognate *onn-ee* 'heart, courage' and Konso *otan-ta* 'center' would still require explanation. The reconstruction of an old EC root **huzn-*, **huzan-* 'heart' might help address those issues and is well supported by the superscripted initial in the Dhasanach cognate *^huođin-ni* 'heart'.⁶⁶ Alternatively, the analysis could fall back on the parsimonious hypothesis of PEC ***hu** yielding **wV** and **øV** in EC with an irregular **w** in Burji.

The lack of a reflex of initial ***h** before the back high vowel ***u** in HEC is a possible counterargument against positing a new phoneme. Still, it is justifiable to argue that (a) the reflexes of the phoneme ***h** in word-initial position are not restricted to the environment before ***u**, as shown in [519] and [698], and (b) PHEC ***hu** might have merged with ***ɦu** and resulted in Burji **hu**, NHEC **wV**.

⁶⁶ The assimilation of the initial fricative ***h** to the second fricative **z** would also accommodate the Arbore cognate *zazzá* 'heart' as more plausible than the assimilation of a glide **w**.

This is not the only merger – an earlier **ħa* merging into **ha* would account for examples like the following:

PHEC [426] **ħaraar y-s-*, **araar y-s-* ‘to mediate, reconcile’ → Burji (*h*)*araar-s-* ‘mediate, reconcile’; PNHEC **ħaraar y-s-*, **araar y-s-* ‘to mediate, reconcile’ → Had *araar-*, *araass-* ‘to reconcile, relieve (from danger)’, Kam *araar-s-* ‘mediate, intervene’, Sid *araar-s-* ‘mediate, reconcile’, Ged *araar-s-* ‘mediate, reconcile’ (see Oro *araarsa* ‘mediate, reconcile’).

The reconstruction of initial **ħ* in [426] is tentatively based on the free variation of initial *h* ~ *ø* in Burji (*h*)*araar-s-* ‘mediate, reconcile’. This makes the alternative reconstruction **araar-s* ‘to mediate, reconcile’ also reasonable by arguing for an analysis of areal epenthesis to explain the variation in Burji.⁶⁷ The lack of **ħa* and the frequent attestation of the PHEC sequence **ħa* (§2.1.5.2) also support the assumption of a merger of **ħa* with **ha* at an early stage, making [426] an exception.

PHEC	Merged	Burji	Had	Kam	Sid	Ged	Environment
	<i>*ħ</i> →	<i>h/ø</i>	<i>h/ø</i>	<i>h/ø</i>	<i>h/ø</i>	<i>h/ø</i>	#_a
<i>*ħ</i> and <i>*h</i> →	<i>*ħ</i> →	<i>h</i>	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>	#_u

The *h* ~ *ø* variation in the reflexes of the merger before *a* may be a later areal phenomenon and includes the reflexes of **ħ*.

As mentioned above, two competing analyses present themselves. The first treats Burji *h* and PNHEC *w/y* as epenthetic segments inserted before certain vowel-initial stems. The epenthesis account would derive *w/y* from hiatus resolution. This analysis has precedents in synchronic Kambaata/Sidaama morphophonology (§3.4.1.2). However, three factors challenge mere epenthesis:

- (a) Distribution asymmetry – epenthesis typically resolves specific morpheme boundaries, but the *h/w* reflexes appear in the roots.

⁶⁷ The variation between *h* and *ø* in initial position has been reported for Burji (Tsfaye Baye, 2015, p. 78) and Kambaata (Treis, 2008, p. 27). Treis (2008, pp. 27, 32) states that initial *h* ~ *ʔ* variation, e.g., Kambaata *ħank’afu* ~ *ʔank’afu* ‘to embrace’ is widespread in East Cushitic and Ethiosemitic languages. However, she stresses that there is no general lack of distinction between *h* and *ʔ* word-initially. Such areal initial epenthetic *h-* is more prevalent in Gedeo and is even attested in established vowel-initial stems, e.g., pronouns.

- (b) External cognates such as Somali **ħ** in cognate items suggest an inherited segment rather than independent innovation.
- (c) Divergent Burji/NHEC reflexes weaken the analysis as epenthesis, since it would predict parallel outcomes; the regular **h/w** split suggests proto-segment conditioning.

The epenthesis mechanism may explain some examples (particularly those lacking external cognates), but cannot account for the full distribution.

The second analysis reconstructs a proto-phoneme ***ħ** with the reflexes presented above. Several methodological considerations favor this analysis.⁶⁸

First, the principle of economy favors the reconstruction of a single proto-phoneme over the positing of two independent epenthesis processes (one in Burji, another in PNHEC) that happen to align perfectly across the attested cognate sets. Second, the comparative method's uniqueness principle holds that a unique correspondence pattern should be analyzed as reflecting a unique proto-phoneme unless there is compelling evidence to the contrary. Third, the EC parallels show a mostly regular correspondence with **ħ** in those languages that retain the pharyngeal. Fourth, the PNHEC **w ~ y** variation (as in Kambaata [519]) is best understood as phonetically natural assimilation of ***ħ** to the following vowel, and alternations in Sidaama *hom-* ~ *wom-* 'be full' suggest an underlying segment rather than epenthetic instability. This is what we expect from an inherited segment undergoing conditioned sound change. Fifth, the epenthesis hypothesis faces the empirical challenge that word-initial epenthesis contradicts the otherwise inter-morphemic position of all other HEC epenthetic processes.

Additional support comes from the vowel quality variation observed across the PNHEC and EC languages in these roots. The alternation between Hadiyya, Sidaama, Gedeo *woʔm-* and Kambaata *wiim-* 'be(come) full' in [520], and more generally the varied vowel qualities in the PNHEC reflexes, may reflect coloring effects from an earlier voiced glottal fricative that lowered or backed adjacent vowels before and after merging with the glides. Furthermore, such vowel variation is common in other EC languages (e.g., WOT) when they occur adjacent to the pharyngeals.

⁶⁸ See exceptions such as Somali [256] *ħan* 'slander, gossip' and [229] *ħanfafaq* 'whisper', which are reconstructable with regular **h** reflexes from PHEC **ham ʔ* 'to slander' and **haf~haf-* 'to whisper', respectively.

The following list evaluates the reflexes of PHEC ***h** attested in different environments.

	Position	Burji	Had	Kam	Sid	Ged	Environment	Evaluation
1.	Initial	h	w	w	w	w	/__* u , (* i)	secure
2.	Initial	w/ø	h	h	h	h	/__* o	tentative
3.	Initial	h ~ ø	ø	ø	ø	ø	/__* a	tentative
4.	Non-initial	y/ø	h	h	ø	y/ø	/V__	reasonable

The observed tendencies include: (1) regular NHEC **w** and Burji **h** before ***u** (and maybe ***i**), (2) sporadic NHEC **h** and Burji **ø/w** before ***o**, (3) tentative NHEC **ø** and Burji free variants **h ~ ø** word-initially before ***a**, following the **h** and **h** merger,⁶⁹ and (4) sporadic NHEC **h**, **ø** and Burji **ø** in non-initial/post-vocalic position (Burji and Gedeo **y** might be reflexes, but the hiatus that a zero reflex creates is also resolvable by epenthetic **-y-/w-**, see §§3.4.1.2, 3.3.3).

Series (1) is evaluated as ‘secure’ based on six supporting examples with consistent reflexes across all five HEC languages. Series (2) and (3) are ‘tentative’ due to limited attestation (two and one examples respectively) and irregular reflexes in individual languages (Burji [467] *wonfa* ‘nine’ showing **w** rather than the expected **h** before **o**). Series (4) achieves ‘reasonable’ status despite variation (Burji/Gedeo **y/ø** vs. Kambaata/Hadiyya **h**) because the contrast with complete loss of ***h** in non-initial position distinguishes ***h** from ***h** diachronically.

In sum, PHEC ***h** is arguably reconstructable on the basis of a regular correspondence set showing Burji **h/ø**: PNHEC **h/w/ø**. While synchronically unattested, its reconstruction is motivated by principles of economy, EC correspondences, and the comparative method. This analysis brings the reconstruction of HEC into closer alignment with standard comparative methodology, eliminating the need to posit word-initial epenthesis processes that are otherwise unattested in the language family. The reconstruction raises the PHEC consonant inventory from 20 to 21 segments.

⁶⁹ Before the low vowel **a**, the reflexes are in free variation **h ~ ø** in Burji (supported by the dictionary listing in Getahun Addo (2017)) and mostly **ø** in NHEC with sporadic **h** in some lexemes.

2.1.5.4 Summary

The velar, pharyngeal, and glottal consonants bring together two themes that recur throughout this chapter: the asymmetry between initial and non-initial positions, and the fact that several “phonological” outcomes are only intelligible once morphophonological context is taken into account.

The reconstruction reveals a contrast between stable initial retention and significant non-initial weakening. The velar stops ***g** and ***k** are securely reconstructed, though they exhibit divergent developments: ***g** is susceptible to palatalization in Burji, while ***k** is frequently lost or lenited in non-initial position unless protected by gemination or clustering. Non-pulmonic ***k'** presents a more complex picture. It is reconstructable for PNHEC but shows signs of marginal status or reintroduction through contact in the modern languages. Among the glottal consonants, ***h** is stable initially but prone to loss elsewhere, and the comparative evidence supports reconstructing ***ʔ** as part of the stem in a subset of items alongside its multiple grammatical functions (genitive/reflexive/nominal and epenthetic uses). The arguable reconstruction of the pharyngeal ***ħ** is supported by distinctive internal and EC correspondences. Together, these patterns signify the role of both internal lenition processes and external contact influence in shaping the HEC phonological inventory.

Finally, this section has already foreshadowed two subgroup-relevant developments that will be relevant again in §2.3: the relative instability of certain onsets in the northern branch of HEC, and the way glottal-related alternations feed into later restructuring.

2.2 HEC Vowels

The HEC languages have five vowel qualities, each contrasting between single and double quantity.

Table 2: *Highland East Cushitic Vowels Inventory*

	Front		Back
High	i, ii		u, uu
Mid	e, ee		o, oo
Low		a, aa	

Following a moraic analysis, I refer to **single V** and **double VV** (for identical-vowel sequences) rather than “short” vs. “long” vowels. All HEC languages

have five single and double vowels (Hudson, 1976, p. 247). Section 2.2.1 treats regular reflexes; Section 2.2.2 addresses deviations in quantity and quality.

2.2.1 Regular Stem Vowels

The sections below present the regular reflexes of stem single vowels across the five HEC languages. They are all securely reconstructed for PHEC and PNHEC. Section 2.2.2 treats deviations in quantity or quality separately, since many of them reflect later changes rather than inherited vowel contrasts.

2.2.1.1 Single Vowels

There are five reconstructable HEC single vowels: **i*, **e*, **a*, **o*, **u*. Some reflexes of the single vowels exhibit irregularities in the corresponding vowel quantities depending on the language.

PHEC **i*

	[1]		[54]		[702]	
PHEC	*ill-	'eye'	*min-	'house'	*hig-	'to return, pass'
Burji	illa	'eye'	mina	'house'	hig-	'exchange'
PNHEC	*ill-	'eye'	*min-	'house'	*hig-	'to return, pass'
Had	ille	'eye'	mine	'house'	hig-	'pass'
Kam	illi-ta	'eye'	mini	'house'	hig-	'pass'
Sid	?ille	'eye(s)'	mine	'house'	hig-	'return'
Ged	ille	'eye(s)'	mine	'house'	hig-	'return, repeat'

Single **i* is also found in sets such as PHEC **[151]** **is-e* → PNHEC **ise* 'her', PHEC **[7]** **is-u* → PNHEC **is-u* 'he', PHEC **[46]** **ik'k-*, **ikk-* → PNHEC **ink'-*, **ink-* 'tooth', PHEC **[237]** **firb* → PNHEC **sirb-* 'to dance and sing', PHEC **[287]** **t'ib-* (?) → PNHEC **t'ib-* 'to make trouble', PHEC **[392]** **lik'-* → PNHEC **lik'-* 'to lend', and PHEC **[637]** **sim-* → PNHEC **sim-* 'ensete tree 1st stage'. However, it is also attested in some series of irregular correspondences discussed in §2.2.2.

PHEC *e

PHEC	[184] *reñ-	'die'	[407] *ebel-, *ebal	'so-and-so'	[223] *eg-	'to wait'
Burji	rey-	'die'	ebeloo	'so-and-so'	eeg-	'to wait'
PNHEC	*reñ ?	'die'	*ebel-, *ebal-	'so-and-so'	*eger	'to wait'
Had	leh-	'die'	ebaro	'so-and-so'	eger-, egar-	'wait (for)'
Kam	reh-	'die'	ebalo	'so-and-so'	eger-, agar-	'wait (for)'
Sid	re-, re?-	'die'	?ebelo	'so-and-so'	?agar-	'wait, watch'
Ged	rey-, re(?)	'die'	ebelo	'so-and-so'	egert-, eged'-	'wait!, wait (for)'

The disyllabic examples above show that **e** and **a** can sporadically assimilate, such as [223] in all NHEC and [407] in Gedeo and Sidaama, or dissimilate, as in Hadiyya and Kambaata [407] (← **ebel-*) and [223] (← **eger*).⁷⁰ It is more common for disyllabic stems to have identical vowels, which supports the dissimilation proposal (consistent with the East Cushitic echo-vowel epenthesis (Sasse, 1979)).

Additional examples include PHEC [114] **debe?* → PNHEC **debe?* 'to belch', PNHEC [75] **t'e?m-* 'to taste good', PHEC [144] **me?* → PNHEC **me?* 'how many? how much?', PHEC [288] **k'e/-* → PNHEC **k'e/-* 'snail', PHEC [514] **ebb-* → PNHEC **ebb-* 'to bring'. Overall, single ***e** is found in limited comparanda in HEC.

Further examples show some irregularity in the quantity of the corresponding vowels due to later morphophonological processes (§2.2.2.2). For instance, PHEC verb [184] **reñ ?* 'die' above is derived from the same stem as the noun PHEC [329] **reñ-f-*, the PNHEC **reñ-f-* 'corpse'; however, other than Kambaata *re/fa* 'corpse', all HEC noun cognates have reflexes with double **ee**: Burji *reesi* 'corpse', Had *lee/fa* 'corpse', Sid *reeffa* 'corpse', Ged *reen/fa* 'corpse' (see Oro *reefa* 'corpse'). The vowel doubling is attributed to the loss of the glottal **h** in that noun.

⁷⁰ PNHEC [223] **eger* correspondences include the unexplainable **-er** which has no parallel in the rest of EC (see PEC **?eeg-* 'watch, look out' (Sasse, 1982, p. 67)).

For examples of an alternation between **e** and **a** before **ʔ**, see PHEC [269] *t'eʔ- 'to diminish', *t'aʔ- 'to extinguish' (same set with alternating **e** ~ **a**) and PHEC [30] *kaʔ- → PNHEC *kaʔ-, *keʔ- 'to get up, arise' (§2.2.2.1).

PHEC *a

	[137]		[2]		[189]	
PHEC	*lal-	'cattle'	*an-i	'I'	*k'araar-	'be bitter'
Burji	lál-i	'cattle'	ʔáni	'I'	k'araar-	'be bitter'
PNHEC	*lal-	'cattle'	*ani	'I'	*k'araar-	'be bitter'
Had	laro	'cattle'	ʔáni	'I'	k'araar-	'be bitter'
Kam	lalu	'cattle'	án	'I'	k'araar-	'be bitter'
Sid	laló	'cattle'	ʔani	'I'	k'araar-	'be bitter'
Ged	lalo	'cattle'	ani	'I'	k'araar-	'be bitter'

Further examples of single PHEC *a include: PHEC [80] *gal- → PNHEC *gal- 'to enter', PHEC [9] *arrab- → PNHEC *arrab- 'tongue', PHEC [14] *az- → PNHEC *az- 'milk', PHEC [36] *gam- 'to bite' → PNHEC *gam-ʔ- → *gaʔm- 'to bite', PHEC [67] *ang- → PNHEC *ang- 'hand, arm', PHEC [68] *malañ- → PNHEC *malañ- 'pus', PHEC [72] *lam- → PNHEC *lam- 'two', PHEC [110] *bar- → PNHEC *barr- 'day', PHEC [137] *lal- → PNHEC *lal- 'cattle', and PHEC [449] *k'al- → PNHEC *k'al- 'give birth, beget (of animals)'.

Gedeo additionally has a reflex **e** of PHEC *a. Such a change of *a to **e** is observed in suffixes such as **-em** (pass) from PHEC *-am (§4.2.1) and **-ed'** (mid) from Oromo **-ad'** (§4.2.3). An example of the passive suffix is Gedeo [545] *dun-em*- 'leak (of e.g. water)'. Wedekind (1980, pp. 139–149) relates the correspondence between **a** and **e** to accent placement in Gedeo (§2.4), where the final syllable of the stem is stressed, while most suffixes and some nominal and verbal forms require penultimate stress.

	PNHEC		Gedeo	
[545]	*dun-am-	'leak (of water)'	dun-em-	'leak (of water)'
[90]	*malab-	'honey'	malebo	'honey'
[444]	*bagaz-	'spear'	bagedo, bagado	'spear'
[707]	*magan-	'sky, God'	mageno, magano	'God, sky'
[59]	*agan-	'month, moon'	agen-jo	'month, moon'
[10]	*hat't'if y	'to sneeze'	het't'if-at-	'sneeze'

In the above examples, Gedeo **e** is found in the penultimate syllable, except in example [10]. The variation between **e** and **a** in [444] and [707] is a difference in the source's transcription: [444] *bagedo* (Wedekind, 2008, p. 74),

bagado (Hudson, 1989, p. 139), [707] *mageno* (Wedekind, 2008, p. 35), *mageno*, *magano* (Hudson, 1989, p. 71).

As Sasse (1982, p. 16) notes, the contrast between **a** and **e** is neutralized in closed syllables before **y** in Burji. That results in a difference in transcription depending on the source, e.g., sometimes Sasse (1982) has **a** where Straube's data has **e**.

PHEC *o

	[478]		[139]		[580]	
PHEC	*kott-	'hoof, foot'	*wozan-	'heart'	*boc'-	'to carve, split'
Burji	kottee	'hoof'	wodana	'heart'	boc'-	'carve'
PNHEC	*kott-	'hoof, foot'	*wozan-	'heart'	*boc'-	'to carve, split'
Had	konte?e	'hoof, foot'	wodano	'heart'	boc'-	'carve, to split'
Kam	kotti-ta	'foot step'	wozana	'heart'	boc'-	'snap (to break)'
Sid	kotte	'hoof'	wodana	'heart'	boc'-	'carve, split'
Ged	kotte	'hoof'	(onna)	'heart'	boc'-	'carve, split'

Further examples of single PHEC ***o** include: PHEC [126] **gog-* → PNHEC **gog-* 'skin, hide', PHEC [87] **k'ot-* → PNHEC **k'ot-* 'to dig', PHEC [95] *fomb-* 'lung', and PHEC [286] **c'ob-* → PNHEC **c'ob-ʔ-* → **c'op'p'y-* 'to drip, to leak'.

PHEC *u

	[48]		[300]		[108]	
PHEC	*tum-	'to pound'	*k'ub-	'finger(-ring)'	*sunk'	'to kiss'
Burji	tum-	'to pound'	k'ubee	'finger-ring'	sunk'-	'kiss'
PNHEC	*tum-	'to pound'	*k'ub-	'finger(-ring)'	*sunk'	'to kiss'
Had	tum-	'pound (metal)'	k'uba?a	'finger-ring'	sunk'-	'kiss'
Kam	tum-	'pound (metal)'	k'ubbata	'finger-ring'	sunk'-	'kiss'
Sid	tum-	'hit, strike'	k'ubb-e	'ring'	sunk'-	'kiss'
Ged	tum-	'pound (metal)'	k'ube	'ring'	sunk'-	'kiss'

The meanings of [300] **k'ub-* include 'finger(-ring)' and 'ring'.

Further examples of single PHEC ***u** include: PHEC [40] **tuf-* → PNHEC **tuf-* 'to spit', PHEC [672] **dul-* → PNHEC **dul-* 'to skin/slaughter', PHEC [180] **utub-* → PNHEC **utub-* 'pole, center-pole of house', PHEC [408] **mut-* → PNHEC **mut-* 'awl for basketwork', PHEC [268] **mur-* → PNHEC **mur-* 'to cut', PHEC [224] **t'ur-* → PNHEC **t'ur-* 'dirt of body, clothing', PHEC [391] **kum-* → PNHEC **kum-* 'thousand', and PHEC [381] **k'ul-*, **k'ul~k'ul~l-ʔ-* → PNHEC **ul-*, **ul~ul~l-ʔ-* 'leech'.

2.2.1.2 Double Vowels

PHEC shows evidence for double-vowel counterparts to every single vowel (*ii, *ee, *aa, *oo, *uu). They are secure reconstructions, though these are less frequent in the current database than single vowels.

PHEC *ii

	[45]		[400]		[250]	
PHEC	*fiik' ?	'to whistle'	*giir-	'fire'	*t'iib-	'to push'
Burji	fiisk'-	'whistle'	jiira	'fire'	d'iib-	'to push'
PNHEC	*fiik' ?	'to whistle'	*giir-	'fire'	*t'iib-	'to push'
Had	fiink'ees-	'whistle'	giira	'fire'	t'iib-	'to push'
Kam	fiik'-	'whistle'	giira-ta	'fire'	t'iib-	'press'
Sid	fiik'a	'whistle'	giira	'fire'	t'iibi	'push'
Ged	fiiss-	'whistle'	giira	'fire'	d'iib-ema	'be pushed'

Additional examples of double PHEC *ii include: PHEC [390] *hiit'* → PNHEC *hiit'* 'to stretch', PHEC [348] **c'ii?* 'bird' → PNHEC **c'ii?-i-ff-*, PHEC [493] **hii?* → PNHEC **hii?* 'take!', PHEC [273] **biin-* 'mosquito, fly (sp.)' → PNHEC **biin-* 'mosquito, fly (sp.)', PHEC [652] **kiil-* → PNHEC **kiil-* 'to do magic', and PNHEC [115] **k'iid-* 'become cold'.

One set remains difficult: [583] is reconstructed as PHEC/PNHEC **maaz ?* alongside **miiz ?* 'to hurt'. The distribution is split, with *ii in Burji *miid'* 'to harm, hurt', Sidaama *miic'i* 'hurt', and Gedeo *miid'a* 'hurt', but *aa in Hadiyya *mada?* 'to wound somebody' and Kambaata *maaz-*, *maz-ees-* 'be wounded, wound somebody'. This could reflect (a) dialect mixture at the PNHEC stage, (b) analogical remodeling in one sub-branch, or (c) contact influence, as suggested by the Oromo parallel *miid'a* 'hurt'.

PHEC *ee

	[462]		[613]		[136]	
PHEC	*meet-	'scraping board'	*t'een-	'rain'	*eel-	'pond, pool'
Burji	meeta	'scraping board'	c'eena	'rain'	eela	'water-hole'
PNHEC	*meet-	'scraping board'	*t'een-	'rain'	*eel-	'pond, pool'
Had	meeta	'scraping board'	t'eena	'rain'	leera	'pond, pool'
Kam	meeta	'scraping board'	t'eena	'rain'	eela	'pond, pool'
Sid	meeta	'scraping board'	t'eena	'rain'	?eela	'low place, pond'
Ged	meeta	'scraping board'	t'eena	'rain'	eela	'puddle, water'

Further examples of ***ee** are mostly reconstructable for NHEC: PHEC [326] **geeg-eess-* → PNHEC **geeg-eess-* ‘accompany’ (see Oromo *geggeessa* ‘accompany’), PHEC [79] **eey-* → PNHEC **ee*, **yee* ‘yes’, PNHEC [42] **k’eer-* ‘long, tall’, PNHEC [406] **wees-* ‘ensete tree (*Ensete ventricosum*)’, PNHEC [460] **beet-* ‘son, child’, PNHEC [741] **beet-*, **beet-beet-* ‘grandchild’, and PNHEC [437] **seer-* ‘law’.

Similar to single ***e**, there are examples of double ***ee** with some irregularity in quantity and sometimes in the quality of the corresponding vowels, e.g., PHEC [168] **heeʔ-* ‘to exist, to live’ is transcribed with double vowels in Hadiyya *heeʔ-* ‘live, exist’, Sidaama *heed’-* ‘exist, be, be in place, be alive’, and Burji *heettaa* ‘here it is!’, but a single vowel in Kambaata *heʔ-* ‘live’ and Gedeo *hed’-* ‘exist, be there’. The following examples suggest a change from PHEC ***ee** to Burji **ay**, but there is no regular substantial evidence to establish such a rule:

	Burji		PNHEC	
[440]	wayne, woynnee	‘Colobus monkey’	*weeʔn-	‘Colobus monkey’
[634]	laymi	‘bamboo’	*leem-	‘bamboo’
[128]	d’aybood’-	‘to become thirsty’	*t’eeb ʔ	‘to be thirsty’

PHEC ***aa**

	[189]		[492]		[209]	
PHEC	*k’araar-	‘be bitter’	*haar-	‘new’	*laaf-	‘soft, weak’
Burji	k’araar-	‘be bitter’	haareyaa	‘new’	laafa	‘soft, weak’
PNHEC	*k’araar-	‘be bitter’	*haar-	‘new’	*laaf-	‘soft, weak’
Had	k’araar-	‘be bitter’	haareécco	‘new’	laafa	‘weary, infirm’
Kam	k’araar-	‘be bitter’	haaroo	‘new’	laafa	‘soft, weak’
Sid	k’araar-	‘be bitter’	haaro	‘new’	laafa	‘weak’
Ged	k’araar-	‘be bitter’	haaro	‘new’	laafa	‘soft, weak’

Similar to single ***a**, double ***aa** is frequently attested in the available database. Here are additional examples: PHEC [635] **maal-* → PNHEC **maal-* ‘meat’, PHEC [252] **hasaaw-* → PNHEC **hasaaw-* ‘speak, talk’, PHEC [258] **maak-*, **ma~mmaak-* → PNHEC **maak-* ‘say in a proverb, tell a story’, PHEC [284] **ayyaan-* → PNHEC **ayyaan-* ‘festival, holiday’ (also found in Ethiosemitic languages), PHEC [354] **yaaʔ-* → PNHEC **yaaʔ-* ‘to flow, assemble’, PHEC [450] **jaal-* → PNHEC **saal-* ‘be shy, be ashamed’, PHEC [548] **arraab-* → PNHEC **arraab-* ‘to lick’, PHEC [703] **kaakk y-s-* → PNHEC **kaakk y-s-* ‘to cackle (hen)’, PHEC [3] **aabb-* → PNHEC **aabb-* ‘father’, PHEC [150] **baall-* → PNHEC **baall-* ‘feather’, PHEC [426] **haraar y-s-*, **araar y-s-* → PNHEC **haraar y-s-*, **araar y-s-* ‘to mediate, reconcile’, and PHEC [282] **aar* ‘be angry’.

Interestingly, it is mostly attested adjacent to the liquids **l** and **r**, and to a lesser extent, the other sonorants, but is least common between obstruents.

PNHEC *oo

	[330]		[8]		[675]	
PHEC	*sooz-	'dawn'	*fool-	'four'	*k'oom-	'calabash'
Burji	sood-	'dawn'	foola	'four'	k'oomoo	'calabash'
PNHEC	*sooz-	'dawn'	*fool-	'four'	*k'oom-	'calabash'
Had	sood-	'dawn'	sooro	'four, lot'	k'ooma	'calabash, gourd'
Kam	sooz-	'dawn'	foolu	'four'	k'oomu	'calabash'
Sid	sood-	'dawn'	foole	'four'	k'oon-co	'calabash, skull'
Ged	sood-	'dawn'	foole	'four'	k'oon-jo	'gourd, vessel'

Additional examples of double ***oo** include: PHEC **[31]** ***oo y**, ***oo ?** 'to cry' → PNHEC ***oo y-**, ***oo ?** 'to cry',⁷¹ PNHEC **[217]** ***fool-** 'breath, odor, smell', PHEC **[610]** ***roob-** → PNHEC ***roob-** 'hippopotamus', PHEC **[603]** ***boos-** → PNHEC ***boos-** 'fireplace', PHEC **[623]** ***oot-** → PNHEC ***oot-** 'fence', PNHEC **[347]** ***oos-** 'boy'.

Gedeo has a few words transcribed with a single **o** that correspond to double **oo** in the rest of NHEC. One example is the regular **oo** in **[675]** **k'oon-jo** 'gourd, vessel, calabash' above, which has a variant transcribed with a single **o** as seen below, in addition to the other two examples.

	PNHEC	Gloss	Gedeo	Gloss
[675]	*k'oom-	'calabash'	k'omma	'gourd, vessel, calabash'
[222]	*t'ook-	'to burst'	d'oʔ-	'burst, let burst; burst'
[327]	*koor-	'saddle'	koricca	'saddle'

Burji and wider EC parallels point to a single **o** for **[222]** Burji **t'okk-** 'to leak, drop, burst', Oromo **d'oʔa** 'burst, explode', Afar **ḏokʕ-** 'bruise, dent', and in **[327]** Burji **koraa** 'saddle', Konso **kor-aa** 'saddle', Saho, and Afar **kor** 'saddle'. These two sets are therefore reconstructed with a single vowel at PHEC level, i.e., **[222]** ***t'ok-** 'to burst' and **[327]** ***kor-** 'saddle'. The short vowel in Gedeo need not, however, be inherited at PNHEC level. Given the recurrent Oromo influence on Gedeo, as argued elsewhere, a secondary short vowel is especially plausible in **[222]**, which has a close Oromo parallel. I therefore treat the Gedeo corresponding vowel here as a local, likely contact-influenced reflex rather than as evidence for a regular PNHEC shortening rule. The

⁷¹ Compound from PEC ***ʔoo y-** 'say oo' (Sasse, 1982, p. 156).

PNHEC long vowels are retained on the basis of the non-Gedeo NHEC evidence.

PHEC *uu

	[351]		[438]		[325]	
PHEC	*t'uut' -	'to suck'	*zuug-	'scrape, tan'	*duud-	'mute (person)'
Burji	t'uunt' -	'absorb'	duug-	'scrape, tan'	duuda	'dumb'
PNHEC	*t'uut' -	'to suck'	*zuug-	'scrape, tan'	*duud-	'mute (person)'
Had	t'uut' -	'to suck'	duug-	'scrape, tan'	duuda	'mute'
Kam	t'uut' -	'to suck'	zuug-	'scrape, tan'	duuda	'mute'
Sid	t'uut' -	'to suck'	duug-	'scrape, tan'	duuda	'mute'
Ged	t'uut' -	'to suck'	duug-	'scrape, tan'	duuda	'blunt, dull'

Double *uu is less frequently attested than the other vowels. Additional examples include PHEC [\[130\]](#) *tuul- → PNHEC *tuul-, *tuur-m- 'to pile up', PNHEC [\[49\]](#) *k'uub-, *k'uub-ʔ- 'egg', PNHEC [\[411\]](#) *buur- 'butter', and PNHEC [\[742\]](#) *buur-, *buur-ʔ- 'to anoint/smear, anoint oneself'.

Two examples show differences in the quantity of corresponding vowels:

PHEC [\[66\]](#) *anuun- 'breast' → Burji *ununa* 'breast'; PNHEC *anuun- 'breast' → Had *anuuna* 'breast', Kam *anuuna* / *unuuna* 'breast', Sid *unuuná* 'breast', Ged *unuuna* 'breast'.

PHEC [\[186\]](#) *c'ub- 'dip in, immerse' → Burji *c'uub-* 'immerse, to dip, soak'; PNHEC *c'ub-, *c'ub ʔ, *c'ub y-s 'to rinse, immerse' → Had *c'ufc'ufaaʔ-* 'to wash, to immerse', Kam *c'uf aʔ* 'rinse, immerse', Sid *c'uuʔi* 'baptize', Ged *c'uup'a* 'dip (in)'.

In [\[66\]](#), the double vowels are continued in PNHEC, but are transcribed with a single **u** in Burji. The initial vocalic alternation of **a** with **u** is an irregular assimilation, as the free variation of the two Kambaata variants suggests. The second example [\[186\]](#) is reconstructed with a single ***u** vowel, but it has a double **uu** in Sidaama, Gedeo, and Burji. These double vowels could be due to contact with Oromo (see *c'uup'a*, *c'up'*-, *c'ub-* 'dip, baptize') or internal alternation processes.

2.2.2 Variation in Stem Vowels

Several comparative sets show irregular correspondences in vowel quality and quantity. Where possible, I group these sets by conditioning environment, since at least some of the variation plausibly reflects later developments rather

than inherited contrasts. The likely sources of irregularity include adjacent consonants (especially glottals and labials), accent placement, morphological boundaries (including remnants of older compounds), vowel ablaut, contact influence, sporadic assimilation, and transcription differences.

2.2.2.1 Irregularity in Vowel Quality

The Influence of the Glottal Stop ʔ

The glottal stop ʔ influences the quality of the preceding vowel. The examples below illustrate variation in vowel reflexes across multiple languages for stems containing ʔ. In each set, the vowel reflex for each language is marked with (x) in the appropriate column. The first three examples are [431] *raʔ- ‘ripen, be ripe’, [30] *kaʔ- (PNHEC *kaʔ-, *keʔ-) ‘to get up, arise’, and [34] *baʔ- ‘be lost’:

	[431]	a	e	i	[30]	a	e	i	[34]	a	e	i
PHEC	*raʔ-				*kaʔ-				*baʔ-			
Burji		x				x					x	
Had				x			x					x
Kam			x				x				x	
Sid		x	x			x					x	
Ged		x				x					x	

For each of the above sets, there are three retentions of PHEC *a. The reflexes of Gedeo and Sidaama are the most stable across the three examples. Burji and Kambaata each reflect a regular a once when compared to the EC parallels. Hadiyya shows fronting to i in all three examples. Kambaata and Sidaama show e in the first two sets. Such diversity of reflexes, including three different vowels, suggests that the conditioning effect of ʔ was not uniform across HEC.

The second set includes [293] *waʔ- ‘water’, [269] *tʰaʔ-, *tʰeʔ- (PNHEC *tʰaʔ-, *tʰeʔ-) ‘be extinguished, to diminish’, and [144] *meʔ- ‘how many? how much?’:

	[293]	a	o	[269]	a	e	[144]	e	i
PHEC	*waʔ-			*tʰaʔ-, *tʰeʔ-			*meʔ-		
Burji		x			x	x			x
Had			x		x			x	
Kam			x		x	x			
Sid		x				x		x	
Ged		x				x		x	

The alternations are limited to only two vowels in these examples. Again, Sidaama and Gedeo show uniformity throughout, while the rest show variation between **a**, **o**, **e**, and **i**. In [293], the rounded **o** in Hadiyya and Kambaata is likely due to the preceding labial **w**, with further incentives from the variation-triggering **ʔ**.

According to the distribution patterns, the attested vowels are chiefly the low vowel **a** and the front vowels **i** and **e**. Burji predominantly uses **a** (in 4 stems), with some **e** (2 stems), indicating a preference for low/mid vowels. Hadiyya uniquely prefers the high front vowel **i** (in 3 stems), with some **e**, **a**, and **o** (1 stem for each). Comparable to Hadiyya, Kambaata also has varied vowel reflexes, using **a** (2 stems) and **e** (3 stems), with **o** (1 stem). Sidaama uses both **a** (4 stems) and **e** (3 stems) reasonably equally, with doublets appearing in roots like **raʔ-* (both **a** and **e** forms). Gedeo consistently uses **a** (4 stems) and **e** (2 stems).

The double vowels, such as in PHEC [749] **yaaʔ-* ‘assembly, meeting’, PHEC [354] **yaaʔ-* ‘to flow, assemble’, and PHEC [168] **heeʔ-* ‘to exist, to live’, are not influenced by the glottal stop. However, exceptions to this generalization can be found in [440] **weeʔn-* ‘Colobus monkey’. Overall, the variation of a single vowel may serve as additional supporting evidence for reconstructing a root or stem with a final glottal stop, in conjunction with other evidence.⁷² Other factors to consider are assimilation in Hadiyya and the influence of the inserted **-y-** in Burji, which contribute to this variation. Additionally, there is a merger of the reflexes of PEC ***ɕ** with **ʔ** in HEC, where the earlier form is also known to condition vowel changes in some EC languages, e.g., Western Omo-Tana. For instance, the parallel of PHEC [431] in Afar is *laʕallaʕe* ‘cook, fry, grill’, *alaye* ‘become ripe, ripen, cook’, and the PEC reconstructions of [30] is **kaɕ-* ‘get up, arise’ and [117] is **faɕ-* / **saɕ* ‘cow’.

The only existing study on Cushitic that has proposed glottal stop influence is Hetzron (1980, p. 107). He discusses its phonetic effect on the following vowel in pronominal prefixes.

⁷² A counterexample would be the variation found in the loanword: PHEC [341] **loom-* ‘lemon’ → Burji *loomay*; PNHEC **loom-* ‘lemon’ → Had *loomeʔe*, Kam *loommi-ta*, Sid, Ged *loome* (see Oro *loomii* and Amharic *lomi* ‘lemon’) unless one accepts that it has been borrowed at the proto-level and used to have a glottal stop or another elided final consonant. The same applies to [365] *bärbäre* ‘ground pepper’ from Amharic.

The Influence of the Labials *b* and *w*

Vowels before the labials **b** and **w** also exhibit some irregular variation in quality. The reflexes are listed below based on whether the vowel reflex is rounded or unrounded:

	[4]	rounded	unround.		[74]	rounded	unround.
PHEC		*gulib-	*gilub-	'knee'	*k'ub-	-	'fall (down)'
Burji	-		(gilba)	'knee'	-	-	-
Had		gurubbo	-	'knee'	ub-	-	'fall (down)'
Kam		gulubi-	-	'knee'	ub-	-	'fall (down)'
Sid		gulube	-	'knee'	ub-	-	'fall (down)'
Ged		gulubo	-	'knee'	-	ib-	'fall (down)'

	[495]	rounded	unround.	
PHEC	-	*lab-		'many, big, much'
Burji	-	laboo		'many, much'
Had		lob	-	'many, big, important'
Kam	-		(ligiba)	'many'
Sid		lobo	-	'big, many, loud, much'
Ged		(lumolla)	-	'big'

	[466]	rounded	unround.	
PHEC			*diwh-, *diwk-	'empty'
Burji		duwwa	-	'empty and deserted house'
Had	-		diiha	'empty'
Kam	-		diiha	'empty'
Ged		duwwi	-	'empty, bare, naked'

The following examples are from NHEC only with an initial **w** reflex of ***h** before ***u**:

	[519] rounded	unround.		[163] rounded	unround.	
PNHEC	-	*hilill-	'smoke'	*hur-	-	'young ox'
Had	wuriira	wiriira	'smoke'	woʔla	-	'calf'
Kam	-	wiliili-	'smoke'	worjaamu	-	'young bull'
Sid	-	wiliile	'smoke'	wor-aamo	war-aamo	'ox, bull'
Ged	-	wiliille	'smoke'	-	-	-

	[607] rounded	unround.		[520] rounded	unround.	
PNHEC	*huym-s-	-	'to fill'	*huy ʔ-m-	-	'be full'
Had	wonf-	wanf-	'fill'	woʔm-	-	'be full'
Kam	-	wiinf-	'fill'	-	wiim-	'be full'
Sid	wonf-	-	'fill'	woʔm-	-	'be full'
Ged	wonf-	wanf-	'fill'	woʔm-	-	'be full'

Reflexes of these two stems are discussed in §4.2.5.7.

There is an example with double unrounded vowel reflexes, but where Gedeo still shows irregular reflexes before the labial **-b** in PNHEC [306] **y ebb-* 'be(come) hot/warm' → Hadiyya, Kambaata, and Sidaama *iibb-*, Gedeo *eebb-* 'be(come) hot/warm'. Below are more examples with single regular vowels without any variation in the reflexes before the labial **-b**:

	[768] *hab-	'to forget'	[183] *gib-	'to refuse, hate'	[300] *k'ub-	'finger(-ring)'
PHEC	*hab-	'to forget'	*gib-	'to refuse, hate'	*k'ub-	'finger(-ring)'
Burji	hab-	'forget'	jibb-	'to hate'	k'ubee	'finger-ring'
PNHEC	*hab-	'to forget'	*gib-	'to refuse'	*k'ub-	'finger(-ring)'
Had	-	-	gib-	'flee'	k'ubaʔa	'finger-ring'
Kam	hab-	'forget'	gib-	'refuse'	k'ubbata	'finger-ring'
Sid	hab-	'forget'	gib-	'refuse'	k'ubb-e	'ring'
Ged	-	-	gib-	'refuse'	k'ube	'ring'

Therefore, the preceding examples show that labials do not exert a regular rounding effect on vowels in comparable environments. Forms such as **hab-* 'to forget' and **gib-* 'to refuse, hate' retain unrounded vowels before **b**, so rounded reflexes in other sets are better treated as sporadic rather than as the expected outcome of labial influence. Such sporadic rounding is attested in Burji and Gedeo [466], and Hadiyya and Sidaama [495]. In [4], however, PNHEC **gulub-* does not allow a confident choice between PHEC **gulib-* and **gilub-* 'knee'.

Some variation in vowel quantity and quality could be attributed to vowel coalescence as a result of intervocalic consonant loss ([115] and [114]) or compounding ([257]).

	[115]		[114]		[257]	
PHEC	*k'ab y-t	'be cold'	*debeʔ-	'to belch'	*ke	'you' (sg.obj)
Burji	kabbee	'be cold'	deemmeys-	'to belch'	jeé	'you' (sg.obj)
	d'aa-					
PNHEC	*k'iid-	'be cold'	*debeʔ-	'to belch'	*ke	'you' (sg.obj)
Had	k'iid-	'be cold'	deeʔacc-	'belch'	keése	'you' (sg.obj)
Kam	gid-	'be cold'	debeeʔl-	'belch'	kees	'you' (sg.obj)
Sid	k'iid-	'be cold'	debeeʔl-	'belch'	-	-
Ged	k'iid-	'be cold'	(deeʔefat-)	'belch'	-	-

The consonant loss and coalescence evident in the verbalized PNHEC stem [115] **k'iid-* 'be cold' occurred via labial lenition and glide loss **k'aid* ← **k'awid* (§§2.1.3, 2.1.4). In [257], Kambaata and Hadiyya combine **ke* and **ise* into **keese*.

Compounding or Affixation

Prior studies attribute the fluctuation of the root vowel to fossilized ablaut in the prefix conjugation in East Cushitic (Sasse, 1982, pp. 42, 26, 30; Sasse, 1979, p. 5). Hetzron (1980, p. 39) also believes that some vowels in the prefix serve as tense markers in East Cushitic languages. He also cites Diakonoff (1965, p. 84), where the latter implies that the vocalic alternation originally occurred within the root, and it has been shifted to the prefix in Cushitic. Taking the previous wider perspective, one might want to see if the remnants of the old prefix or verb compounding system may have played any role in the verb-initial vowel ablaut in HEC.

	[306]		[514]		[223]	
PHEC	*y ebb-	'be hot/warm'	*ebb-	'to bring'	*eg-	'to wait'
Burji	d'aybad'	'to be warm'	amb-	'bring'	eeg-	'to wait'
PNHEC	*y ebb-	'be hot/warm'	*ebb-	'to bring'	*eger	'to wait'
Had	iibb-	'be hot/warm'	eeb-	'bring'	eger-, egar-	'wait (for)'
Kam	iibb-	'be hot/warm'	eeb-	'bring'	eger-, agar-	'wait (for)'
Sid	iibb-	'be hot/warm'	?abbi	'bring'	?agar-	'wait'
Ged	eebb-	'be hot/warm'	-	-	egert-, eged'-	'wait!, wait (for)'

As [306] and [514] demonstrate, potential labial influence is evident in these examples. However, remnants of support verbs might be a more plausible

explanation within the context of the current research. At least [306] consists of the support verbs *y 'to say, do, be' (§4.2.4) attached to a following (nominal) stem.⁷³ Example [514], on the other hand, might be a remnant of a prefix verb similar to PHEC [62] *-a/- 'give birth, beget (of human)' and [449] *k'a/- 'give birth, beget (of animals)'. As mentioned previously (§2.2.1.1, PHEC *e), [223] could also be alternatively explained in terms of vowel assimilation (e → a in V₁) if one discards the prefixing possibility.

Numerals

The vowels and consonants of the numerals are mostly unstable, sometimes with more than one variant recorded in the same source. This is expected because they are the most susceptible to use in contact situations, and their variants are comparable to those found in other languages in the area.

	[38]		[19]		[142]	
PHEC	*liñ-	'six'	*sazhet	'eight'	*tam-n-	'ten'
Burji	liya, lia	'six'	hiditta	'eight'	tanna	'ten'
PNHEC	*leñ-	'six'	*sazzeett-	'eight'	*tom-n-	'ten'
Had	loho	'six'	sadeento	'eight'	tommo	'ten'
Kam	lehu	'six'	hezzeettu	'eight'	torduma	'ten'
Sid	lee, leye	'six'	sette	'eight'	tonne	'ten'
Ged	-	-	saddeeta	'eight'	tomme	'ten'

Similar to [142], Burji has a different vowel in PHEC [636] *omt-, *om~omt- 'five' → Burji *umutta* 'five'; PNHEC *ont- 'five' → Hadiyya *onto* 'five', Kambaata *ontu* 'five', Sidaama *onte*, *ʔoonte* 'five', and Gedeo *onde* 'five'.

Some of the above examples include one irregular reflex in a single language, but there are more complex cases.

PHEC [51] *lokk-, *lekk- 'foot, leg' → Burji *lukka* 'foot, leg'; PNHEC *lokk-, *lekk- 'foot, leg' → Had *lokko* 'foot, leg', Kam *lokka-ta*, Sid *lekka*, Ged *lekka* 'foot, leg'

PHEC [206] *awd- 'threshing floor' → Burji *oyda*, *ooyda* 'threshing floor'; PNHEC *awd- 'threshing floor' → Had *awda*, *oodoʔo*, Kam *oodu*, Sid *ʔaiddumme*, *ayddumme*, Ged *oodde* 'threshing floor'

⁷³ The support verb is placed before the verb as a remnant of the older order (i.e., when affixes used to occur before the root, so too did the support verbs, e.g., Gedeo *mar-go* → *imar-* 'take' and *dag-* 'come' → *idag-* 'bring').

Three different corresponding vowel reflexes (**u**, **o**, **e**) are found in [51], and the vowel and glide sequences **oy**, **ooy**, **ai**, **aw**, and **oo** occur in [206]. Therefore, two proto-forms are reconstructed for [51] while the **u** of the Burji form is attributed to Oromo influence since the latter has *luka* ‘foot, leg’. [206] **awd-* ‘threshing floor’ is more problematic, and thus is only provisionally reconstructed.

According to previous studies, vowel variations have been reported to include vowel laxing in closed syllables and word-internally in the HEC languages, with the exclusion of Sidaama. This laxing most commonly affects **a** and **i**, less commonly **e**, and rarely **o** and **u**. The lax allophones are **ə**, **ɪ**, **ɛ**, **ɔ**, and **ʊ**, respectively (Hudson, 1976, p. 249). For Burji, Sasse (1982, p. 16) describes these as vowel allophones that apply in cases of both single and double vowels. He also notes that the contrast between **a** and **e** is neutralized in closed syllables before **y** (see §2.2.1.2, PHEC **ee*).

2.2.2.2 Irregularity in Vowel Quantity

There are several examples with irregularity in the quantity of vowels (double vowels corresponding to single ones), and in a few cases, quality as well. Where related derived stems exist, one may tentatively assume remnants or innovated morphophonological processes triggered by derivational morphology and analogy.

The first example includes the words PHEC [184] **reñ* ? ‘to die’, [439] **reñ-* ‘death’, and [329] **reñ-f-* ‘corpse’, which share the same root originally.

	Verb		Noun		Noun	
Burji	rey- , re-	‘die’	reya , reeya , reeha	‘death’	reesi , reefa	‘corpse’
Had	leh-	‘die’	leho	‘death’	leeja , leejja	‘corpse’
Kam	reh-	‘die’	rehu-ta	‘death’	re(e)ja	‘corpse’
Sid	re- , re?-	‘die’	reo , rewo	‘death’	reejja , reeja	‘corpse’
Ged	rey- , re(?) -	‘die’	reyo , reo	‘death’	reenja , renja	‘corpse’

The parentheses around the Kambaata and Gedeo sounds indicate optional segments, e.g., in Kambaata, variants with single and double vowels have been recorded for both the noun and the verb.

Other sets derived from a common root include PNHEC [338] **yaʔ-* 'to call',⁷⁴ and PHEC [749] **yaaʔ-* 'assembly, meeting'.

	Verb		Noun	
Burji	-	-	yaa	'assembly, meeting'
Had	yaaʔa, yaʔ-	'to call'	yaʔ-	'large in mass'
Kam	yaaʔ-	'hold (a meeting)'	yaaʔa	'assembly of elders'
Sid	yaaʔa	'come together'	yaaʔa	'communal assembly'
Ged	yeʔa	'to call'	yaaʔa	'crowd'

The following is reconstructable for PHEC [168] **heeʔ-* 'to exist' and **heeʔ-f-* 'life, existence' (see §4.1.2):

	Verb		Noun	
Had	heeʔ-	'live, exist'	heecca	'life, condition'
Kam	heʔ-	'live'	heecca-ta	'life, existence'
Sid	heed'-	'exist, be alive'	heeffo	'life'
Ged	hed'-	'exist, be there, live'	heeffo, hed'o	'life'

Some irregularity in quantity in the above examples could be due to the loss of the following glottal, for instance, in the derived nouns in **reñ-f-* 'corpse' and **heeʔ-f-* 'life, existence'.

Variation in vowel quantity is also attested in the PHEC verb [548] **arraab-* 'to lick' and noun [9] **arrab-* 'tongue'.

	Verb		Noun	
Burji	arrab-, arraab-	'lick'	arraaba	'tongue'
Had	aall-	'lick'	allabo	'tongue'
Kam	arraab-	'lick'	arrabi-ta	'tongue'
Sid	-	-	ʔarrabo	'tongue'
Ged	arraab-, arraap'p'-	'lick'	arrabo	'tongue'

A more complex connection is proposed between the derived verbs PHEC [306] **y ebb-* → **d'ayb-* 'be(come) hot/warm' and [128] **t'eeb ʔ* 'to be thirsty' and the noun **t'eeb-* 'thirst'.

⁷⁴ Set [354] with Burji *yaʔ-* 'to flow' and Gedeo *yaaʔa* 'flow' could also be argued to be cognates.

130 A Lexical and Phonological Reconstruction of Highland East Cushitic

	Verb		Verb		Noun	
Burji	d'aybad'	'be hot/warm'	d'aybood'-	'become thirsty,	d'ayboo	'thirst'
Had	iibb-	'be hot/warm'	t'eebe?-	'be thirsty'	t'eebo	'thirst'
Kam	iibb-	'be hot/warm'	t'eebek-, t'eebe?-	'be very thirsty'	t'eebe?u	'thirst'
Sid	iibb-	'be hot/warm'	-	-	-	-
Ged	eebb-	'be hot/warm'	d'eeboo?-, d'eebot-	'be thirsty'	d'eebo	'thirst'

A less clear semantic link could be assumed for PHEC [287] **t'ib-* (?) 'to make trouble', [250] **t'iib-* 'to push', and [417] **t'ib-* 'disease, evil'.

	Verb		Verb		Noun	
Burji	d'ib-	'to make trouble'	d'iib-	'to push'	d'iba	'evil'
Had	t'ib-	'to hide'	t'iib-	'to push'	t'isso	'demon, evil'
Kam	-	-	t'iib-	'press'	t'ida-ta	'illness'
Sid	d'ib-	'be troubled'	t'iib-	'to push'	t'ibba	'disease'
Ged	d'ib-	'cause trouble'	d'iib-	'press, push'	d'ibe	'sickness'

One final example with three stems includes PHEC [252] **hasaaw-* (← **hafaaw-*) 'speak, talk', [229] **haf~haf-* 'to whisper', and **haas-* 'speech'.

	Verb		Verb		Noun	
Burji	haasaaw ⁻⁷⁵	‘speak, talk’	hasaas-	‘whisper’	hasaasaa	‘whisper’
Had	hasaaw-	‘chat’	hajafa	‘whisper’	(wac’a)	‘speech’
			wac-			
Kam	haasaaww-	‘discuss’	hajajf-	‘whisper’	haasaawwa	‘speech’
Sid	hasaaw-	‘speak, talk’	hajajf-	‘whisper’	-	-
Ged	haaso?	‘speak, chat’	hajaf-	‘whisper’	haaso?a	‘speech’

The above examples demonstrate that variation in vowel quantity can sometimes be the only difference between nouns and verbs, or derived verbs, with the verb having a single vowel against a double vowel in the noun, or vice versa, with or without additional suffixes. The other factor that might have contributed to the alternation in quantity is accent shift between the ultimate and penultimate syllables, since it is governed by syllable weight (§2.2). Therefore, this depends on the attached suffix and the subsequent degree of apocope that occurred in each word.

⁷⁵ The first syllable has a single vowel in *hasaawa* 'to talk' as transcribed in Getahun Addo (2017, p. 426). The dictionary transcribes the two verbs and the noun with a single vowel.

Here are pairs of nouns that are clearly semantically comparable, which could have led to the irregularity in their vowel correspondences through analogy:

	Noun		Noun	
Burji	meena	'people, man, person'	meena	'people'
Had	mán-cq	'man, person'	meent-iccq, mento	'woman'
Kam	man-cu	'man, woman'	meent-iccu-ta	'woman'
Sid	man-co	'man, woman'	man-co	'man, woman'
Ged	manjo, manna	'man, person'	manjo, manjicco	'woman'

They are reconstructed as PHEC [138] **man-f-* (pl) 'person' and [215] **meen-t-* 'woman, girl' (see [609] **man~n-* 'people'). Derivation has contributed to the variation, and paradigmatic drift is responsible for the identical lexemes.

	Noun		Noun	
Burji	c'iid'aa, c'iid'd'aa	'bird'	c'iw?e, c'iw?eena	'chicken'
Had	c'ii?-icco	'small bird'	c'iic'oolla,	'chick'
Kam	c'ii?-iccu-ta, c'iiccu-ta	'bird'	c'iic'oorra	'chick'
Sid	c'e?icco, c'ee?-icco	'bird'	c'aac'c'urre	'chick'
Ged	c'iicco, c'i?a	'bird'	c'aac'ute	'chicken'

The two nouns above are reconstructed as PHEC [348] **c'ii?-* 'bird' and [226] **c'iw?-* 'chick(en)'. According to Gasparini (1983, p. 56), Sidaama *c'e?icco* is contractable to *c'icco* 'bird'. Anbessa Teferra (2000, p. 27) explains that vowel deletion occurs when two different vowels occur in a sequence in Sidaama. However, Gasparini transcribes the remaining vowel *i* with a falling accent, indicating that its underlying quantity is double.

The above categorization should be taken as an arguable and provisional approximation with possible explanations. The variation in vowel quality and quantity requires further research to establish potential regularities or rules.

2.2.3 Terminal Vowels

Highland East Cushitic nominal stems end in vowels, referred to here as terminal vowels (TV). This section examines the distribution and variation of terminal vowels across the five HEC languages. Section 2.2.3.1 discusses identical TVs where cognates share the same final vowel quality. Section 2.2.3.2 examines cases of TV variation where cognate sets show different vowel qualities.

According to Hudson (1976, p. 249), HEC (Burji (Tsfaye Baye, 2015, p. 59) and Gedeo (Eyob Kelemework, 2015, p. 267)) words end in vowels except in Hadiyya and in a few instances in Kambaata. Such vowels in final position, i.e., terminal vowels (TV), are only reconstructed on the pronominals as arguably fossilized nominal suffixes (§4.1.5.2). The terminal vowels of the nouns may have originated from fossilized suffixes, recent independent suffixation, apocope of stem-final consonants, or coalescence with other suffixes.⁷⁶ Other than the arguably fossilized nominal suffixes, the alternative origins fall outside the scope of this analysis. Thus, terminal vowels are hyphenated, suggesting a morpheme boundary following the root final consonant.

Terminal vowels are widely reported as devoiced or extra-short in Burji (Sasse, 1982, p. 17; Tsfaye Baye, 2015, p. 46), Hadiyya (Tadesse Sibamo, 2015, p. 27), Kambaata (Treis, 2008, p. 20), and Gedeo (Eyob Kelemework, 2015, p. 42), while Sidaama appears more resistant to this reduction. Hudson (1976, p. 250) notes that devoicing is especially common after strong (or heavy) syllables and may render terminal *i* and *a* nearly inaudible; he also observes that final vowels may be lost altogether in continuous speech in Hadiyya, and less commonly in Kambaata.

2.2.3.1 Identical Terminal Vowels

The HEC languages show instances of identical corresponding terminal vowels. Only three vowels, *a*, *e*, and *o*, are commonly attested as terminal vowels. As terminal vowels, *i* and, to a lesser extent, *u* are only found on pronouns.

Terminal vowels are chiefly single vowels. However, there are instances of double terminal vowels, some of which are due to differences in transcription. For example, where Roba Dame & Wedekind (2008) have double terminal vowels, Hudson (1989) has a single terminal vowel in Burji data. The history of stress placement (§2.4) and contact has also contributed to such an anomaly.

⁷⁶ Underived verb roots typically end in consonants. Overall, the morphophonemic and morphological evidence shows that the PHEC roots must have ended in consonant(s), and those with vowels are their stems (including apocopated ones).

Terminal -i

Terminal **-i** is found on pronominals only.

	[2]		[27]		[6]	
PHEC	*an-i	'I' (subj)	*at-i	'you' (sg.subj)	*is-i	'she' (subj)
Burji	ʔáni	'I' (subj)	afi	'you' (sg.subj)	ʔífi	'she' (subj)
PNHEC	*ani	'I' (subj)	*ati	'you' (sg.subj)	*isi	'she' (subj)
Had	ʔáni	'I' (subj)	ʔáti	'you' (sg.subj)	ʔísi	'she' (subj)
Kam	án	'I' (subj)	át	'you' (sg.subj)	íse	'she' (subj)
Sid	ʔani	'I' (subj)	ʔáti	'you' (sg.subj)	ise	'she' (subj)
Ged	ani	'I' (subj)	ati	'you' (sg.subj)	ise	'she' (subj)

For sets [2] and [27], Hudson (1989) transcribes Kambaata *ani* 'I' and *ati* 'you' with the terminal vowels, while *án* 'I' and *át* 'you' from Treis (2008) have devoiced and inaudible terminal vowels. In [6], the irregular terminal vowel of Kambaata, Sidaama, and Gedeo *ise* 'she' is probably due to analogy with the accusative form of the pronoun *ise* 'her' (obj).

The demonstratives provide additional evidence of these terminal vowels (§4.1.5.1).

Terminal -e

Terminal **-e** is the least attested final vowel in comparison to final **-a** and **-o**. This section includes most of the examples where terminal **-e** is attested in the available data.

	[151]		[28]		[240]	
PHEC	*is-e	'her'	*ay-e	'who?'	*k'ort'umm-	'fish'
Burji	ʔifeé	'her (obj)'	ayyee	'who?'	k'urt'ummee	'fish'
PNHEC	*ise	'her'	*aye	'who?'	*k'ort'umm-	'fish'
Had	ʔise	'her (obj)'	ayye	'who?'	k'urt'umeʔe	'fish'
Kam	isé-ta	'her (obj)'	aye	'who?'	k'urt'ummee-ta	'fish'
Sid	ʔisé	'her (obj)'	aye	'who?'	k'ult'uʔme	'fish'
Ged	isee	'her (obj)'	ayye	'who?'	k'ult'uʔme	'fish'

The terminal vowel **-e** is only arguably reconstructed for the pronominals, as a fossilized suffix. It is a feminine accusative ***-e** in [151] **is-e* 'her' and [28] **ay-e* 'who?'. The nominative counterparts are [6] **is-i* 'she' and [28] **ay-i* 'who?' (→ Kambaata and Sidaama *ayi* (nom) 'who?').

Further examples with identical terminal **-e** attested in the HEC languages are [273] **biin-* ‘mosquito, fly (sp.)’, [478] **kott-* ‘hoof, foot’, [91] **mooyy-* ‘mortar’, [169] **k’unc’-* ‘bark (of tree)’, [264] **arr-* ‘grey hair’, and [440] **weeʔn-* ‘Colobus monkey’.

Some instances are only identical in NHEC, excluding Kambaata:

PNHEC [1] **ill-* ‘eye’ → Had, Sid, Ged *ille* ‘eye’ (see Kam *illi-ta* ‘eye, seed, grain’ and Burji *illa* ‘eye’).

PNHEC [544] **c’ub-* ‘dagger, knife’ → Had *c’ubeʔe*, Sid *c’uube*, Ged *c’uup’e* (see Kam *c’up’i-ta* ‘dagger’).

PNHEC [271] **ar-* ‘wife’ → Had, Sid, Ged *are* ‘wife’, (see Kam *ari-ta* ‘wife’, PNHEC [212] **ar-o* ‘husband’).

As shown in the above examples, nouns are always reconstructed without the terminal vowel.

Some loanwords have identical terminal **-e** vowels. for instance, [240] **k’ort’umm-* ‘fish’ from Oromo *k’urt’ummii* ‘fish’ (Sasse, 1982, p. 129) and other sets listed above. More loanwords are found with the terminal vowel **-a**.

Terminal -a

The terminal vowel **-a** is the most frequent in HEC, and the following are a few examples:

	[67]		[68]		[167]	
PHEC	<i>*ang-</i>	‘hand, arm’	<i>*malaḥ-</i>	‘pus’	<i>*fayy-</i>	‘healthy’
Burji	<i>anga,</i> <i>ganaa</i>	‘palm, hand’	<i>malaa</i>	‘pus’	<i>fayya-n-ga</i>	‘healthy’
PNHEC	<i>*ang-</i>	‘hand, arm’	<i>*malaḥ-</i>	‘pus’	<i>*fayy-</i>	‘healthy’
Had	<i>anga</i>	‘hand, arm’	<i>mara</i>	‘pus’	<i>fayyaʔa</i>	‘healthy’
Kam	<i>anga-ta</i>	‘hand, arm’	<i>mala-ta</i>	‘pus’	<i>fayya</i>	‘healthy’
Sid	<i>ʔanga</i>	‘hand, arm’	<i>malaa</i>	‘pus’	<i>fayyaaleessa</i>	‘healthy’
Ged	<i>anga</i>	‘hand, arm’	<i>mala</i>	‘pus’	<i>fayya</i>	‘healthy’

The above examples show the addition of suffixes following the terminal vowel: **-ta** in Kambaata, **-n-ga** in Burji, and **-a** with epenthetic **-ʔ** in Hadiyya (§3.3.4). The reflexes of PHEC [68] **malaḥ-* ‘pus’ indicate that some identical terminal vowels are a result of apocope.

Further examples with identical terminal **-a** vowels are found in the individual languages on the following PHEC stems: [66] **anuun-* ‘breast’, [126] **gog-* ‘skin, hide’, [383] **orz-* ‘hair of animals’, [209] **laaf-* ‘soft, weak’, [284] **ayyaan-* ‘festival, holiday’, [462] **meet-* ‘scraping board for use in preparing ensete’, [547] **am-* ‘mother’, and [609] **man~n-* ‘people’.

Some instances of terminal **-a** might be part of (fossilized) plurative marking (§4.1.1.2) where there is reduplication **C_i~C_i-a*, e.g., PHEC [609] **man~n-* ‘people’ (see [138] **man-f-* ‘person’).

Similar to **-e**, numerous loanwords have identical terminal **-a** vowels. At least in Kambaata, Treis (p.c.) notes that the consonant-final Amharic loanwords receive a default terminal **-a**.⁷⁷

Unknown source: Burji [349] *gabalaa* ‘limit’, Had, Kam, Sid, Ged *gabala* ‘border, limit’.

Amharic *atār* ‘peas’: Burji [283] *ataraa* ‘peas’, Had, Kam, Sid, Ged *atara* ‘peas’.

Amharic *fəmbəra* ‘chick-peas’: Burji, Had, Sid, Ged [611] *jimbura* ‘chick-peas’, Kam *jumburu*, *jumbura* ‘chick-peas’ (see Oro *jumburaa* ‘chick-peas’).

The source of [349] is unknown, but it is identified as a loanword because of the irregular **l** in Hadiyya in place of the expected **r** reflex (§2.1.1.3, PHEC *l). The direction of transfer of the last two examples is debatable. Additional evidence of loanword status is the retention of intervocalic **b**, which should have been lenited in intervocalic position (§2.1.3).

Hudson (1989, p. 12) notes that in all five languages, though less often in Sidaama, the unstressed **-a** is pronounced as [ə]. This observation is relevant for the many instances of reduction of vowels in terminal position in HEC, e.g., inaudible Kambaata **-i** in [2] *án* ‘I’ and [27] *át* ‘you’ discussed above. The stress is obviously on **V₁** in the two examples, while the unstressed terminal vowels are reduced to zero (see §4.1.5.2 (Nominative)).

⁷⁷ Two more examples are in sets [325] Amharic *dida*, *dəda* ‘mute’, Oromo *duudaa* ‘mute’ and [350] Amharic *sānga*, Oromo *sangaa* ‘gelding, bullock, ox’.

Terminal -o

Identical terminal **-o** vowels are most frequently attested in NHEC, where Burji has a different vowel. It regularly corresponds to **-u** in Kambaata with a few exceptions. The following are some rare examples, including Burji.

PHEC	^[407] * <i>ebel-</i> , * <i>ebal-</i>	'so-and-so'	^[198] * <i>tum-am-f-</i> , * <i>tum-t-i-ff-</i> , * <i>tum-t-</i>	'blacksmith'	^[707] * <i>magan-</i>	'sky, God'
Burji	ebeloo	'so-and-so'	tuntoo	'blacksmith'	maganoo	'mist'
PNHEC	* <i>ebel-</i> , * <i>ebal-</i>	'so-and-so'	* <i>tum-am-f-</i> , * <i>tum-t-i-ff-</i> , * <i>tum-t-</i>	'blacksmith'	* <i>magan-</i>	'sky, God'
Had	ebaro	'so-and-so'	tumaanco	'blacksmith'	-	-
Kam	ebalo	'so-and-so'	tumaan-cu	'blacksmith'	maganu	'God'
Sid	?ebelo	'so-and-so'	tunto	'blacksmith'	magano	'God, sky'
Ged	ebelo	'so-and-so'	tunto	'blacksmith'	mageno	'God, sky'

In set ^[407], **ebel-*, **ebal-* 'so-and-so' behaves similarly to pronouns with TV **-o** and **-e** in Hadiyya and Kambaata. In other words, they show gender distinction between the masculine accusative Hadiyya *ebaro*, Kambaata *ebalo*, and the feminine accusative Hadiyya *ebare*, Kambaata *ebale* 'so-and-so'. Kambaata **-o** is an irregular reflex, as the expected vowel is **-u**. The masculine counterpart **-o** of the feminine accusative **-e** above is also inferred to be an arguably fossilized suffix (§4.1.5.2) from the pronoun in ^[292]:

PHEC ^[292] **is-o* 'him' → Burji *ʔisj*; PNHEC **is-o* 'him' → Kam *isú*, Sid *ʔisó*, Ged *isoo* (see Had *it't'o* 'him').

Kambaata *isú* 'him' shows a regular **-u**. In Burji, the vowel **-o** of the object pronoun has likely been restructured in analogy with that of the subject pronouns ^[7] *ʔisj* 'he, it' and ^[6] *ʔiffj* 'she' and became *ʔisj* 'him'.

Further examples with identical final **-o** in the HEC languages include: PHEC ^[235] **od-* 'news', ^[477] **hobol-* 'boat', PNHEC ^[90] **malab-* 'honey', ^[70] **afoʔ-* 'mouth, opening, language', ^[14] **az-* 'milk', ^[444] **bagaz-* 'spear', ^[439] **reñ-* 'death', ^[276] **ball-* 'brother-in-law, father-in-law'. Burji's non-identical terminal vowels corresponding to NHEC **-o** in these examples are either **-a** or **-i**.

The HEC singulative suffixes receive a terminal vowel **-o** (§4.1.1.1) or the regular Kambaata reflex **-u**.

Had, Sid ^[65] *faraffo*, Kam *faaffu*, Ged *faracco* 'horse'.

Had [138] *mán-cɔ* ‘man, person’, Kam *man-cu* ‘man’, *man-cu-ta* ‘woman’, Sid *man-co* ‘man, woman, man, person’, Ged *manjo* ‘man, person’.

Burji [498] *woccoo*, Had *waficco*, Kam *wof-iccu*, Sid *wofficco*, Ged *wolco* ‘dog’.

Had [602] *kin-c* ‘stone, rock’, Sid *kin-co* ‘stone, rock’, Ged *kin-jo* ‘stone, rock’.

Had [602] *kin-c* ‘stone, rock’ is another instance of terminal vowel reduction, which is quite common in Hadiyya.

Concerning Kambaata **-u**, there are examples where Treis (2008) and Treis (p.c.) have TV **-u** while Hudson (1989) sometimes transcribes TV **-o** and vice versa.

	Treis (2008; p.c.)	Hudson (1989)	
[230]	t’abaru	t’abaro	‘ashes’
[144]	meʔu	meʔo	‘how many? how much?’

Such transcription variation does not exist in the TV **-o** of the singulative, which is consistently transcribed as **-u**.

Finally, there are no correspondences with identical terminal **-u** in HEC. Terminal **-u** is mainly found in Kambaata, corresponding to **-o** in the rest of HEC, as mentioned above. Treis (p.c.) notes that there are individual instances of alternating **-u** ~ **-i** variants in Kambaata nouns.

2.2.3.2 Variation in Terminal Vowels

There are several reflexes that show variation in vowel quality across the individual HEC languages. The following examples include all HEC languages, with none having identical TV, although some of the regular correspondences discussed above occur.

138 A Lexical and Phonological Reconstruction of Highland East Cushitic

	PHEC	Gloss	Had	Ged	Sid	Kam	Burji
[174]	*hamas-	'snake'	-affa	-me	-o	-u	-i
[137]	*lal-	'cattle'	-o	-o, -i	-o	-u	-i
[444]	*bagaz-	'spear'	-o	-o	-o	-u	-i
[95]	*fomb-	'lung'	-	-o	-o	-	-i
[606]	*gubeed-	'thigh'	-o	-a	-a	-	-i
[9]	*arrab-	'tongue'	-o	-o	-o	-i-ta	-a
[4]	*gulib-	'knee'	*-co	-o	-co, -e	-i-ta	-a
[382]	*oll-	'neighbour'	-aʔa	-aʔa	-a	-oo	-a
[51]	*lokk-, *lekk-	'foot, leg'	-o	-a	-a	-a-ta	-a
[293]	*waʔ-	'water'	-o	-aʔe, -eʔe	-a	-a	-a
[88]	*gudum-	'shoulder'	-o	-o	-	-u	-a
[195]	*fal-, *falfal-	'omen, magic'	-o	-o	-	-a	-aa
[5]	*san-	'nose'	-e	-o	-o	-u-ta	-a
[439]	*reñ-	'death'	-o	-o	-o	-u-ta	-a
[206]	*awd-	'threshing floor'	-a	-e	-umme	-u	-a
[408]	*mut-	'awl for basketwork'	-a	-a	-C'C'e	-a	-a
[381]	*k'ul-, *k'ul~k'ul~l-ʔ-	'leech'	-a	-a	-e	-a	-ee
[1]	*ill-	'eye'	-e	-e	-e	-i-ta	-a
[415]	*afal-	'liver'	-e	-	-e	-i-ta	-a
[54]	*min-	'house'	-e	-e	-e	-i	-a
[550]	*ibb-	'louse'	-a	-e	-e, -iicco	-i-ta	-a
[224]	*t'ur-	'dirt of body, clothing'	-a	-e	-e	-i	-ee
[544]	*c'ub- (PNHEC)	'dagger'	-eʔe	-e	-e, -a	-i-ta	-

The examples show how the terminal vowel varies. A dash “-” indicates that no cognate is attested in the available sources. Kambaata **-ta** is the feminine accusative suffix, and it was likely added after the TV (mostly **-i** and **-u**). Some of the examples include further suffixes, e.g., the singulative or derivation suffix in [174], [550], and [4]⁷⁸. In [408], Sidaama **-C'C'** includes a glottalized stem-final **-t't'** probably triggered by a glottal suffix (§3.3.1). In the examples where there is more than one TV for the same stem and the same language, e.g., Sidaama **-e**, **-a** in [544] and Gedeo **-o**, **-i** in [137], each is from a different source. They most likely include mistranscriptions, especially in the Gedeo example, because Gedeo nouns do not end in TV **-i**. Kambaata also has variant forms mistranscribed by Hudson (Treis, p.c.).

⁷⁸ Hadiyya [4] **-co** has assimilated to the stem-final **-b**, yielding a geminate in *gurubbo* ‘knee’. This might also have happened to other examples with a terminal **-o**.

Regardless of the variation, if one looks for patterns, Sidaama and Kambaata cognates show the highest number of systematically corresponding TVs, with this occurring in more than half of the above 23 examples.

	Sidaama	Kambaata
[224]	-e	-i
[415]	-e	-i-ta
[550]	-e	-i-ta
[1]	-e	-i-ta
[4]	-e	-i-ta
[54]	-e	-i
[174]	-o	-u
[444]	-o	-u
[137]	-o	-u
[206]	-umme	-u
[439]	-o	-u-t
[5]	-o	-u-ta
[293]	-a	-a
[51]	-a	-a-ta

Sidaama **-e**, **-o**, and **-a** correspond to Kambaata **-i**, **-u**, and **-a**, respectively.

The only TV that shows some degree of similarity across all HEC languages is **-a** in the following examples, with variant TVs indicated in parentheses:

	Hadiyya	Gedeo	Sidaama	Kambaata	Burji
[51]	(-o)	-a	-a	-a-ta	-a
[293]	(-o)	-aʔe, (-eʔe)	-a	-a	-a
[382]	-aʔa	-aʔa	-a	(-oo)	-a
[408]	-a	-a	(-C'C'e)	-a	-a
[381]	-a	-a	(-e)	-a	(-ee)

This level of TV **-a** similarity is expected since it is the most common among the identical terminal vowels (§2.2.3.1).

The most frequent terminal vowels in HEC are **-a**, **-o**, and **-e**.⁷⁹ Terminal **-u** is not attested as an identical TV in cognate sets (except demonstratives), and **-i** appears only on pronominals. The data reveal correspondence patterns between Burji and NHEC. Where Burji has **-a**, NHEC languages (Gedeo, Hadiyya, and Sidaama) often have a corresponding **-o**. Kambaata frequently has **-u** or **-i** corresponding to **-o** and **-e**, respectively, in the rest of NHEC. The terminal vowels are removed from the stem when the nominal stem is derived; therefore, they are not part of the root. However, since HEC nouns must end in a vowel, TVs fulfill an obligatory phonotactic requirement. The alternative hypothesis about the pre-fossilization functions of the terminal vowels and a discussion of the above examples of TV variation are detailed in §4.1.5.2.

2.2.4 Summary

PHEC had a five-vowel system with contrastive quantity: ***i**, ***e**, ***a**, ***o**, ***u** and their double counterparts ***ii**, ***ee**, ***aa**, ***oo**, ***uu**. This system is securely reconstructed for PHEC and PNHEC and remains stable across all five HEC languages.

The regular stem vowels show consistent reflexes in cognate sets (§2.2.1). Single vowels are well attested, while double vowels, though less frequent in the database, are equally secure. The vowel ***e** is relatively infrequent compared to other vowels, a pattern that persists across HEC.

Deviations from regular correspondences fall into two categories (§2.2.2). Quality irregularities often result from adjacent consonants, particularly glottal stops, which condition vowel raising or lowering before **ʔ**. Quantity irregularities, where single vowels correspond to double vowels, reflect morphophonological processes, contact influence (especially from Oromo), or differences in transcription across sources.

Terminal vowels on nominal stems also exhibit both systematic correspondences and variation (§2.2.3). The vowels **-a**, **-o**, and **-e** are the most frequent TVs, while **-i** is restricted to pronominals. Two systematic patterns characterize TV variation: (a) Burji **-i** regularly corresponds to NHEC **-o**, and (b) within NHEC, Sidaama **-e** corresponds to Kambaata **-i**, while

⁷⁹ The same order of frequency and absence of final **u** or **i** have been noted synchronically in Hadiyya (Tadesse Sibamo, 2015, p. 27) as has the lack of final **u** in nouns in Burji (Tsfaye Baye, 2015, p. 48).

Sidaama **-o** corresponds to Kambaata **-u**. Terminal vowels are devoiced in all HEC languages except Sidaama.

These findings support the reconstruction of a stable five-vowel system with quantity contrast for PHEC. The documented irregularities in vowel quality, quantity, and terminal position reflect processes such as phonetic conditioning, morphophonology, apocope, and fossilized suffixation (§4.1.5.2), and contact.

2.3 Syllable and Root Structure

This section provides an overview of the syllable structure in PHEC and PNHEC. It focuses specifically on the development of these structures into their reflexes within Burji, Hadiyya, Kambaata, Sidaama, and Gedeo.

I use “open syllable” for syllables ending in a vowel (CV, V) and “closed syllable” for syllables ending in a consonant (CVC, VC). Syllable weight is treated moraically: light syllables (V, CV) contrast with heavy syllables created by vowel doubling and/or codas (CVV, VVC, CVC, CVVC).

The reconstructed PHEC and PNHEC forms suggest a common syllable structure of CV, CVC, and CVVC types. This system is characterized by a preference for simple onsets and codas at the root level, but with notable instances of double vowels, geminate consonants, and limited heterogeneous consonant sequences, mostly due to morpheme boundary alternation processes.

Thus, PHEC syllable structure follows the template (C)V(V)(C), permitting both open and closed syllables. The following subsections examine onsets (§2.3.1), nuclei (§2.3.2), codas (§2.3.3), and common root structures (§2.3.4). Syllable weight and its relationship to accent are discussed in §2.4.

2.3.1 Syllable Onsets

The onset of syllables in both PHEC and PNHEC is a single consonant (C). There are no initial consonant clusters (e.g., CCV). Onsets can be categorized into single consonant and zero onset.

The onsets of most of the reconstructed stems occur with a single consonant, including plosives, fricatives, nasals, and liquids, for example:

	PHEC	Gloss
[40]	*tuf-	'to spit'
[351]	*t'uut'-	'to suck'
[117]	*saʔ-	'cow'
[307]	*mit-	'one'
[209]	*laaf-	'soft, weak'
[304]	*wit'-	'to sow'

Overall, the plosives and non-pulmonic consonants function as common onsets; thus, syllables with low sonority onsets are most frequent.

The seemingly vowel-initial words in the HEC languages often have complex histories that obscure their original onset. In reconstruction, I therefore allow zero onset for a subset of PHEC/PNHEC forms rather than positing a systematic word-initial *ʔ (§2.1.5.1, PHEC *ʔ). One significant sound change defining PNHEC is the loss of the old initial non-pulmonic velar *k' (§2.1.5.1, PHEC *k'). The reconstruction PHEC [29] *k'ab- → PNHEC *ab- 'to hold/seize' is the best example: Where Burji and other EC languages, such as Oromo, retain forms with k', the PNHEC languages reflect a root *ab-, leading to reflexes like Kambaata ap'p'. This loss of a C onset systematically changed the syllable structure of a whole set of words from CVC to VC in the common ancestor of the Northern branch. Additionally, the weakening and loss of initial *h is another major source of vowel-initial words. Examples of vowel-initial words include:

	PHEC	Gloss
[206]	*awd-	'threshing floor'
[298]	*al ʔ	'to be useful'
[136]	*eel-	'pond, pool'
[62]	*al-	'give birth, beget (of human)'

These examples show reflexes with and without consonant-initial onset in EC, demonstrating the variable stability of this onset.

In summary, PHEC onsets consist of single consonants or zero (some vowel-initial forms derive from the loss of k' and h in PNHEC).

2.3.2 Syllable Nuclei

The nucleus of the syllable is a vowel. Both single (V) and double (VV) vowels are phonemic (§2.2.1) and represent the peak of the sonority in the syllable.

Single vowels are common in many reconstructions (§2.2.1.1), e.g.:

	PHEC	Gloss
[268]	*mur-	'to cut'
[547]	*am-	'mother'
[339]	*ful-	'to go out, exit'
[415]	*afal-	'liver'

The most common vowel nuclei are **a**, **u**, and **i**. Terminal vowels (§2.2.3) further affect surface syllable structure by adding final open syllables to otherwise closed roots.

Though not to the extent of single vowels, double vowels are also frequent (§2.2.1.2), creating heavy syllables such as in the following lexemes:

	PHEC	Gloss
[635]	*maal-	'meat'
[390]	*hiit'-	'to stretch'
[330]	*sooz-	'dawn'
[325]	*duud-	'mute (person)'

A small number of vowel+glide sequences and heterovocalic sequences (especially in Burji and Gedeo) are best treated as later developments; relevant examples are discussed in §2.2.1.2, PHEC *ee. Section 2.2.2.2 includes examples of irregularity in vowel quantity and potential involvement of morpheme boundary alternation processes and morphological derivation.

2.3.3 Syllable Codas

Syllables can be open (ending in a vowel, e.g., CV) or closed (ending in a consonant, e.g., CVC). When closed in reconstructed roots, the coda structure shows some variation. The codas include single consonants (C), geminate consonants (CC), and consonant clusters (cC).

Single consonant codas are the most common type of closed syllable at the reconstructed root level, e.g.:

	PHEC	Gloss	Single Codas
[5]	*san-	'nose'	CVC
[36]	*gam-	'to bite'	CVC
[38]	*liḥ-	'six'	CVC
[136]	*eel-	'pond, pool'	VVC
[551]	*it-	'to eat'	VC

Forms with added vowel suffixes (e.g., case) are therefore treated separately. For example, pronominal [151] **is-e* 'her' reflects the root **is-* plus accusative

***-e** (§4.1.5.2). The root itself has the shape VC, but once the suffixal vowel is added, the final consonant is resyllabified as the onset of the following syllable, *i.se*. Such forms are therefore not counted here as closed root syllables.

The geminate (doubled) consonant codas (CC) are less frequent at the reconstructed root level but are still attested, e.g.:

	PHEC	Gloss	Single/geminate Codas
[9]	*ar.rab-	'tongue'	VC.CVC
[703]	*kaakk y-s-	'to cackle (hen)'	CVVCC-VC
[150]	*baall-	'feather'	CVVCC
[167]	*fayy-	'healthy'	CVCC
[447]	*c'ubb-	'sin'	CVCC

However, many of these geminates might be products of earlier boundary alternation processes, discussed in §§3.1.1 and 4.1.1.2 (Reduplication *C_i~C_i-a). Certain cases of assimilation of the pronominal **-t** to the stem-final **C** (§3.1.1) produce geminates – an innovation in NHEC, supporting Burji's earlier split from the other HEC languages.

Root/stem-final consonant clusters (represented as cC) are rare. Such morpheme-final clusters appear particularly before derivational suffixes, as discussed in Chapters 3 (Alternations at Morpheme Boundaries) and 4 (Morphology), e.g., [215] *meen-t- 'woman, girl'. Sometimes the same derivational suffix yields different morphophonological reflexes in individual languages: regressive/progressive assimilation, fusion, palatalization, metathesis, and epenthesis. However, there are a couple of words reconstructed with **cC** clusters where **c** is usually a sonorant, and **C** is an obstruent, such as:

	PHEC	Gloss	Cluster Codas
[67]	*ang-	'hand, arm'	VcC
[85]	*arg-	'to lend/borrow'	VcC
[95]	*jomb-	'lung'	CVcC
[161]	*bark-	'head-support, pillow (of wood)'	CVcC

Two metathesis processes affecting coda consonants are especially important here because they also serve as diagnostic innovations for subgrouping within HEC. Unique to Burji is the metathesis of a velar and a sibilant. The limited examples of this metathesis include the derivational **-s** following stems:

	PHEC	Intermediate stage	Burji	Gloss
[45]	*fiik' ?	*fiik'-s-	fiisk'-	'to whistle'
[701]	*gig-	*gig-s-	jisk-	'to demolish, destroy'

For [701], Burji also has an underived stem *jig-* 'to demolish/destroy'. In Burji, the **Cs** sequence metathesized to *fiisk'-* and *jisk-*, respectively. In the NHEC languages, however, this sequence evolved differently (e.g., Gedeo *fiiss-* 'to whistle', Hadiyya *fiink'ees-* 'whistle', showing assimilation and other changes). This stop-sibilant metathesis is a distinct innovation of the Burji line after its split from what would become PNHEC.

A compelling piece of evidence for the unity of the NHEC languages as a group separate from Burji is their shared treatment of stem-final sonorant plus glottal sequences.

	[285]		[170]		[120]	
PHEC	*balʔ-	'broad, wide'	*sum ?	'urinate'	*fil-ʔ-	'to comb'
Burji	balʔad'-	'become wide, broad'	sinʔ-aaw-	'urinate'	sil-, silʔ-	'comb'
Ged	baʔla	'wide, broad'	ʃiʔn-	'urinate'	fil-, fiʔl	'comb'

The reconstruction for the above examples contains the sequence **cʔ** (see also [22]). Burji preserves this order faithfully with its reflexes **lʔ** and **nʔ**. However, the NHEC languages, such as Gedeo, show reflexes with the metathesized sequences **ʔl** and **ʔn**. These Burji and Gedeo reflexes are the most transparent examples of this shift. Such metathesis (discussed in §3.3.2) represents a shared innovation that must have occurred in the common ancestor of the NHEC languages after the split from Burji, supporting the status of PNHEC as a valid linguistic node.

In general, the complex consonant sequences discussed here are evaluated as part of reconstructed root structure, not as straightforward reflections of later surface word forms. Many clusters and geminates arise through suffixation or later phonological developments presented in the following chapter, and forms that later take a terminal vowel or a vowel suffix undergo resyllabification.

2.3.4 Common Word and Root Structures

Based on the analysis of onsets, nuclei, and codas, the following primary root structures can be identified for monosyllabic and disyllabic roots:

Monosyllabic	Disyllabic
(C)VC	(C)V.CVC
(C)VVC	(C)VC.CVC
(C)VCC	
(C)VcC	

The parentheses around the consonant (C) indicate the presence or absence of an initial consonant, and the period (.) between the disyllabic syllables marks the syllable boundary. The reconstructed roots typically end in consonants, yielding closed final syllables in the underlying structure. The closed final syllables in PHEC roots attract the accent (§2.4), while open syllables are restricted to initial position in disyllabic words or to affixes.⁸⁰

Hudson's (1976) and Wedekind's (1980) designation of **CV** as the most frequent syllable structure reflects synchronic reality, i.e., nominals end in a terminal vowel by default (§2.2.3). This also explains Wedekind's basic syllable structure formula (-C)CV(V)(-), representing eight different syllable types found in Burji, Gedeo, and Sidaama. However, Wedekind's (1980, p. 141) claim that "no closed syllables have been postulated" is an interesting analytical stance. He argues that it simplifies the description of consonant clusters. In this model, a word like Burji *gilb*- 'kneel' is likely syllabified as *gi.lba*, avoiding a closed *gil* syllable. A consonant like the **l** here might be considered "extra-syllabic" or part of a complex onset in the following syllable. However, this is an analytical choice for phonological modeling. For all practical purposes, CVC structures are indeed phonetically and functionally present. The above analysis shows reconstructable sonorant-obstruent clusters and explicitly acknowledges codas. Thus, the phonotactic reality includes word-medial consonants that close preceding syllables.

2.3.5 Summary

In summary, the reconstructed forms of both PHEC and PNHEC point to a phonotactic system that favors simple syllables like **CV** and **CVC**. The system is also rich with phonemic vowel doubling and various types of codas, which create heavy syllables (**CVV**, **CVVC**, **CVC**). Consonant geminates and clusters are avoided in word-initial position but can appear in the coda, especially at morpheme boundaries.

The change of PHEC reflexes in the daughter languages was driven by a series of phonological innovations related to consonant sequences. The

⁸⁰ There are rare cases of roots reconstructed with a single consonant or vowel only, including the support verbs *y 'to say, do, be' and *ʔ 'to do'.

primary split between Burji and the NHEC languages is fully reflected in their divergent rules for consonant clusters, while the unity of the NHEC languages is clearly demonstrated by shared innovations like the sonorant-glottal metathesis (§3.3.2) and loss of initial ***k'** and ***h** (§2.1.5). Within HEC, further micro-variations, such as the direction of assimilation or the specifics of diphthongs in Gedeo, show the ongoing change of the system.

2.4 Accent

In Highland East Cushitic, accent (often realized as greater prominence and/or higher pitch) is most clearly observed in the distribution of prominence across the last two syllables of the word. Based on the comparative evidence available here, I propose that the historically conservative pattern targets the rightmost heavy syllable (VV or closed), with penultimate accent emerging in many modern forms through later reductions affecting final syllables.

In the various HEC languages, the discussion around accent (or stress) placement has been confined to the penultimate and ultimate syllables. It seems that the accent predominantly falls on the penultimate syllable synchronically. As mentioned above and illustrated below, this is likely a result of historical changes, such as the reduction of terminal double vowels and apocope.⁸¹

This section examines accent patterns in Highland East Cushitic languages, with a particular focus on the relationship between syllable weight, vowel quantity, and accent placement. The analysis demonstrates that the accent historically falls on the ultimate heavy syllable (that having a double vowel or being closed by a consonant), with systematic shifts to the penultimate in contexts of terminal vowel reduction. These patterns provide additional evidence for the moraic analysis of HEC vowel quantity.

⁸¹ Data sources: Burji (Tesfaye Baye, 2015, pp. 63–64; Wedekind, 1980, p. 140), Hadiyya (Tadesse Sibamo, 2015, pp. 31–32), Sidaama (Kawachi, 2007, pp. 63–66), Gedeo (Eyob Kelemework, 2015, pp. 91–92; Wedekind, 1980, p. 140), Kambaata (Treis, 2008, pp. 52–53).

	Penultimate		Ultimate	
Burji	k'álaḡ	'child'	sotteé	'bowel'
	máalaḡ	'meat'	koyloó	'cheek'
Hadiyya	wóʔo	'water'	-	-
	maára	'meat'	-	-
Sidaama	búko	'dust'	mundée	'blood'
	seéda	'long, tall'	aróo	'husband'
Gedeo	béltoḡ	'daughter' (subj)	giirá	'fire'
	báabbə	'Mother!' (voc)	biiʔnné	'mosquito'
Kambaata	sáʔu	'cow' (m.nom)	saʔá	'cow' (m.acc)
	jóolit	'four' (f.nom)	tassáa	'heifer' (m.acc)

Hadiyya does not place an accent on the final syllable (Tadesse Sibamo, 2015, pp. 31–32; Dereje Adane, 2022, pp. 23–24) due to the significant reduction of TVs. As a result, the former penultimate syllable can be interpreted as the current final one when the last vowel is inaudible. In their discussions of stress and accent in Gedeo, neither Dawit Tilahun (2019, pp. 53–55) nor Eyob Kelemework (2015, pp. 91–95) provides examples of accentuation on the penultimate syllable of stems that are not extended with a suffix or in citation forms. However, Wedekind (1980, p. 140) does present examples and notes that certain forms of Gedeo nouns that he calls the non-objective cases and the singular imperative of the verb also have penultimate stress, even in the absence of an “affix”.

Unlike Burji and Kambaata, the Hadiyya and Sidaama languages demonstrate a different placement of the accent within the second mora of bimoraic syllables. Kawachi's (2007, p. 64) analysis explains that in Sidaama, “when the penultimate syllable is a heavy syllable, the second vowel component of it is more prominent than the first one”.⁸² This observation likely applies to Hadiyya as well. For Kambaata, Treis (2008, p. 52) places the accent marker on the first mora, even though both morae in dimoraic syllables have equal prominence. This distinction also supports the terminology used in this study, where the terms “single” and “double” are employed to describe the vowels of HEC languages, emphasizing quantity based on morae rather than simply length.

⁸² However, Kawachi (2007, p. 64), in disagreement with all his predecessors, claims that the high pitch accent occurs on the penultimate vowel segment, not on the penultimate syllable, as shown in the penultimate and ultimate examples.

The placement of the accent in HEC languages is often linked to vowel quantity, particularly with the double TVs (§2.2.3). Consequently, even when an accent appears on a single TV, it is historically regarded as a double vowel (Hudson, 1989, p. 6; Hudson, 2006, p. 296). This assumption has also been argued to be true synchronically (Hayward, 1997, p. 113). According to Wedekind (1980, p. 141), long (or double) vowels do not occur in word-final position in Sidaama and Gedeo, unless final accent/stress is analyzed as “length”. For Wedekind, the Sidaama examples with unaccented double terminal vowels should be interpreted as equivalent to accented single vowels. Thus, Kawachi’s (2007, pp. 63–66) Sidaama examples could be interpreted as having either unaccented double vowels or accented single vowels. However, these alternative interpretations are not strictly necessary since vowel quantity and accentuation are not mutually exclusive, at least from a historical perspective. Additionally, Hudson’s general rule (1976, p. 248) states that stress tends to fall on “strong” syllables, providing a broad framework for understanding the weight of stressed syllables in the HEC languages.

However, previous studies (Hudson, 1976, 1989, 2006; Hayward, 1997; Wedekind, 1980; Sasse, 1982) have not taken into account the potential loss of final consonants, leading to exceptions that challenge the rules they establish. In terms of vowel quantity and syllable position, the accent typically falls on the final heavy syllable. In some individual languages, there are exceptional cases where the accent is placed on an open ultimate syllable with a final single vowel (considered a light syllable). These cases may either derive from a historical or underlying presence of double vowels, as previously suggested, or result from apocope, where the final consonant is lost, resulting in an accented open light syllable (see §4.1.4.1, Genitive Vowel and Suprafix and §4.1.4.4).

Although the above examples are confined to the citation forms, the same general rule of accent placement is assumed for the forms extended by the addition of suffixes, at least underlyingly or historically. That does not preclude the possibility that the individual languages have innovated new systems of accent placement while keeping in mind the historical origin of TVs as remnants of fossilized suffixes (§§2.2.3, 4.1.5.2).

The previous discussion shows the correlation between sonority levels and the placement of the accent. In HEC languages, both voiced and voiceless TVs exist. Typically, the voiceless TVs are not accented and may even go untranscribed in some instances, e.g., Kambaata *ísu* ~ *ís* ‘he’ (nom). The

placement of the accent on the nominative and accusative pronouns in HEC serves as a clear example of how sonority influences accent placement.

	Voiced -V	Accusative	Voiceless -V	Nominative
Burji	ʔiʃeé	'her' (acc)	ʔiʃi	'she' (nom)
	ʔisiʃ⁸³	'him' (acc)	ʔisi	'he' (nom)
Hadiyya	ʔise	'her' (acc)	ʔisi	'she' (nom)
	ʔit't'i	'him' (acc)	ʔit't'o	'he' (nom)
Sidaama	isé	'her' (acc)	ise	'she' (nom)
	isó	'him' (acc)	isi	'he' (nom)
Gedeo ⁸⁴	iseé	'her' (acc)	ise	'she' (nom)
	isé-(eʔe)		ise	
	isoó	'him' (acc)	isi	'he' (nom)
	isó-(oʔo)		isi	
Kambaata	isé-ta	'her' (acc)	ise	'she' (nom)
	isú	'him' (acc)	is(u)	'he' (nom)

Some case vowels are accented in HEC. The accusative vowels are found accented in Burji, Sidaama (Kawachi, 2007, pp. 180, 68–69), Gedeo (Eyob Kelemework, 2015, pp. 54, 112; Dawit Tilahun, 2019, p. 137), and Kambaata (Treis, 2008, pp. 108, 79). The genitive case is also associated with accent in Sidaama and Kambaata (§4.1.4). The nominative counterparts of the above forms are not accented because the case vowels are extremely short and devoiced (Wedekind, 1980, p. 137; Tadesse Sibamo, 2015, p. 86). Although the above examples are only the third-person singulars, the same applies to the other persons.

The manifestations of the apocope process are described for various reflexes of TVs as lax, devoiced, and almost inaudible in Burji, Hadiyya, and Kambaata (Hudson, 1976, p. 250), devoiced in Burji, Sidaama, and Gedeo (Wedekind, 1980, p. 137), shortened and devoiced in Kambaata as well as in K'abeena and Alaaba (Treis, 2008, p. 20), more specifically, devoiced, extra-short, barely audible, and tending to be realized as zero in allegro speech in Kambaata (Treis, 2008, p. 20, fn. 6), similarly lost in continuous speech in

⁸³ Tesfaye Baye's (2015, p. 140) transcription of ʔisiʃ 'him' (acc) with devoiced and accented terminal vowel is likely a mistranscription. Hence, it is transcribed without the devoicing symbol here.

⁸⁴ The first-listed Gedeo forms in each example are from Dawit Tilahun (2019), and the second forms with the glottal stop are from Eyob Kelemework (2015). The sources sometimes do not indicate the accent on the form but describe it, so I added them to all forms for clarity.

Hadiyya and less commonly in Kambaata (Hudson, 1976, p. 250), and extremely short and devoiced in Hadiyya (Tadesse Sibamo, 2015, pp. 27, 68, 86).

The quality and quantity of the syllable nucleus are crucial for explaining the observed accent patterns in the two designated syllable positions (penultimate and ultimate) across HEC. For instance, double vowels tend to carry the accent more frequently than single vowels, while voiceless vowels do not receive any accent. The following illustration roughly incorporates these three sonority variables alongside the positional variables, which work together to determine accent placement. The focus is on the accent placement on the ultimate position (terminal vowels) and the relative placement on the penultimate syllable. In the following schema, “**x**” indicates that an accent is possible on that syllable type, while “-” indicates that an accent is not attested.⁸⁵

	Terminal Vowels	Ultimate Syllable	Penultimate Syllable
Sidaama	Double	x	-
	Single voiced	-	-/x
	Single voiceless	-	x
Gedeo	Double	-	x
	Single voiced	-/x	-
	Single voiceless	-	x
Burji	Double	x	-
	Single voiced	-/x	-/x
	Single voiceless	-	x
Kambaata	Double	x	-
	Single voiced	-/x	-/x
	Single voiceless	-	x
Hadiyya	Double	-	x
	Single voiced	-	-/x
	Single voiceless	-	x

Read as a broad tendency rather than a categorical rule, the schema suggests a diachronic drift: systems that still permit accented final VV (as in Sidaama) look more conservative, while systems with reduced or devoiced TVs

⁸⁵ These might not apply to every instance in the individual HEC languages, but they are proposed as a general historical framework in light of the examples discussed above. Grammars should be consulted for language-specific rules.

increasingly shift prominence to the penultimate syllable. Burji and Kambaata appear intermediate in this respect,⁸⁶ and Hadiyya represents the most advanced reduction of TVs, with no accented final vowels in the data considered here. The “-/x” under single voiced TV indicates that the syllable can be accented or not. The conditional relationship shows that there must be one accented “x” syllable (ultimate or penultimate), so there are no instances of the absence of accent on both the ultimate “-” and penultimate “-” syllables or of the presence of accent (high prominence) on both the ultimate “x” and penultimate “x” syllables of the same word.

In cases where there is “-/x” for both single voiced and voiceless TVs in Burji and Kambaata, the presence of “-” in the voiced TV (i.e., ultimate) means “+” (i.e., accented) penultimate and vice versa. Therefore, this schema illustrates a shift in accent placement from the ultimate syllable (TVs) to the penultimate syllable, which clearly occurs in all cases of single voiceless TVs and the cases where TVs are not accented.

PHEC accent placement followed a weight-sensitive pattern, falling on the ultimate heavy syllable (that containing a double vowel or being closed by a consonant) and on the penultimate syllable otherwise. The modern reflexes show systematic variation: Sidaama and Gedeo retain ultimate accent on double TVs; Burji and Kambaata show intermediate patterns with both double and single accented voiced TVs; Hadiyya exhibits the most advanced reduction, with accent consistently on the penultimate syllable due to complete devoicing of TVs. These patterns support the moraic analysis of HEC vowel quantity adopted throughout this study.

2.5 Summary and Conclusions

This section synthesizes the phonological findings presented in §§2.1–2.4 and contextualizes the reconstruction of PHEC within the broader HEC grammatical framework.

The Proto-Highland East Cushitic phonological system includes a five-vowel inventory with contrastive quantity, relatively symmetrical consonant subsystems, and distinctive prosodic patterns. PHEC reconstructions explain

⁸⁶ This analysis differs from Hayward’s (1997) proposal regarding the reconstruction of both double and single terminal vowels, which he proposes even for PEC. However, as previously mentioned, this proposal does not take into account the possibility of apocope resulting in final consonant loss. Additionally, Hayward’s hypothesis imposes the influence of North Omotic before any attempt to provide an internal explanation.

the reflexes across five daughter languages (Burji, Hadiyya, Kambaata, Sidaama, and Gedeo) using comparative evidence from Cushitic cognates, morpheme boundary alternations, and distributional constraints. The consonant inventory shows series in labial, alveolar, post-alveolar, velar, pharyngeal, and glottal positions. Non-pulmonic consonants contrast at three places of articulation: alveolar (**t'**), post-alveolar (**c'**), and velar (**k'**), with obstruent (sibilant and velar) harmony processes further complicating their reflexes. These patterns distinguish inherited features from areal innovations. The study addresses diverse phonological developments by analyzing correspondence patterns, minimal pairs, and alternations at morpheme boundaries.

2.5.1 Consonant System

PHEC had a consonant inventory of 21 phonemes. The obstruents (obst.) and sonorants (son.) constitute two major HEC sound classes.

Table 3: *Proto-Highland East Cushitic Consonants*

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

This proto-inventory is smaller than the 27 consonants collectively attested across the modern HEC languages (**Table 1**), reflecting the exclusion of contact-induced consonants like **d'** and **s'** (§2.1.2). Voiceless **p** is also not reconstructed as a PHEC phoneme; it appears in Burji primarily through loans from Oromo. Plumonic **c** and **j** are not reconstructed as PHEC phonemes; they appear as derived or loaned sounds in individual languages.

The phonological system reveals both stable correspondences across HEC languages and regular changes in specific languages, as summarized in the following paragraphs.

PHEC had voiceless stops (***t, *k, *ʔ**) and voiced stops (***b, *d, *g**) at three places of articulation. Non-pulmonic ***t'**, ***c'**, and ***k'** are also present in three positions. The labial stop **b** shows variation with **β** and **w** in intervocalic

positions in Burji, Sidaama, and Gedeo. Non-pulmonic *c' shows regular correspondence patterns across the HEC languages and has numerous Eastern Cushitic cognates, supporting its status as inherited, though some instances may reflect contact influence. The glottal stop *ʔ is often lost post-vocally, but functions as a suffix in various verbal derivations supporting its phonemic status.⁸⁷

PHEC has three primary fricatives: *f, *s, *ʃ. Unlike *s, the evidence for the reconstruction of *ʃ is more tenuous; its correspondences are sporadic. Most ʃ reflexes suggest a merger with s, though some cases preserve the ʃ. The voiced fricative *z has been maintained only in Kambaata (and Alaaba); otherwise, it merges with d in Burji, Hadiyya, Sidaama, and Gedeo.

Similar to the glottal stop *ʔ, the glottal fricative *h is often lost post-vocally. The pharyngeal fricative *ħ represents an arguable reconstruction based on initial h/w correspondences and non-initial traces. External evidence from East Cushitic cognates showing ħ supports its reconstruction, though its attestation remains limited (§2.1.5.3). The distribution suggests *ħ and *h merged word-initially before *u and *a, creating the frequent PHEC *ħu and *ha sequences, with distinctiveness preserved only in specific environments.⁸⁸

PHEC had two nasals, *m and *n, with stable correspondences across the daughter languages. PHEC *n is rare in initial position except in pronominal forms. The lateral *l and trill *r yield consistent reflexes in HEC, with specific mutual sound changes in Hadiyya.⁸⁹

The glides *w and *y show variable reflexes, often being lost intervocally or word-finally. They also participated in various morphophonological

⁸⁷ The following sections discuss stops and non-pulmonics across the three chapters: Phonology §2.1.1.1, §2.1.2, §2.1.5; Boundary alternations §3.1.1.1, §3.1.1.5, §3.1.2.1, §3.3 (stop assimilation), §3.2 (stops lenition); Morphology §§4.2.3.1, 3.2.2 (verbal derivation).

⁸⁸ Sections relevant to the fricatives across the three chapters include: Phonology §2.1.1.2, §2.1.3, §2.1.5.2, §2.1.5.3; Boundary alternations §3.2.2 (velar lenition), §§3.1.1.3, 3.1.2.2 (fricative assimilations), §3.3.1.4 (fricative glottalization); Morphology §4.2.2 (causative).

⁸⁹ Discussion of nasals and liquids is found in Phonology: §2.1.1.3, §2.1.1.4, §2.1.3; Boundary alternations: §§3.1.1.2, 3.1.1.4 (nasal assimilation), §3.4.2 (nasal insertion), §3.3.2 (nasal metathesis), §3.1.2.3 (nasal palatalization); and Morphology: §4.1.4.1, §4.1.1.2 (noun morphology), §4.2.1 (verb morphology).

processes, e.g., **y* triggered palatalization of consonants, while **w* had rounding effects.⁹⁰

2.5.2 Reflexes of Reconstructed Consonants

The reflexes of reconstructed PHEC phonemes show both regular inheritance patterns and instances of contact-induced change. The lists below summarize these reflexes, separating securely inherited phonemes from those with evidence of significant contact influence. The reconstructions follow the evaluation criteria established in §1.5.1, with each proto-phoneme assigned a confidence rating based on the regularity and breadth of its attestation.

The most reliable correspondences occur with stops, nasals, and most vowels. More complex patterns are observed with sibilants, which are affected by both regular sound changes and areal features like obstruent harmony and contact influence. Detailed correspondences and reflexes are presented below.⁹¹

Table 4: *Reconstructed HEC Consonants and their Correspondences*

PHEC	Burji	PNHEC	Had	Kam	Sid	Ged	Environment	Confidence
<i>*b</i>	b/w	<i>*b/*w</i>	b/w	b/w	b/β/w	b/β/w	#_, V_V	secure
<i>*t</i>	t	<i>*t</i>	t	t	t	t	all	secure
<i>*d</i>	d	<i>*d</i>	d	d	d	d	all	secure
<i>*z</i>	d	<i>*z</i>	d	z	d	d	all	secure
<i>*s</i>	s	<i>*s</i>	s	s	s	s	all	secure
<i>*ʃ</i>	s/ʃ	<i>*ʃ</i>	s	s/ʃ	s/ʃ	s/ʃ	varies	arguable
<i>*f</i>	f	<i>*f</i>	f	f	f	f	all	secure
<i>*l</i>	l	<i>*l</i>	l, r	l	l	l	#_, V_V	secure
<i>*r</i>	r	<i>*r</i>	l, r	r	r	r	#_, V_V	secure
<i>*m</i>	m	<i>*m</i>	m	m	m	m	all	secure
<i>*n</i>	n	<i>*n</i>	n	n	n	n	all	secure
<i>*w</i>	w/ø	<i>*w/*ø</i>	w/ø	w/ø	w/ø	w/ø	varies	secure
<i>*y</i>	y/ø	<i>*y/*ø</i>	y/ø	y/ø	y/ø	y/ø	varies	secure

⁹⁰ The following are all the glide-related sections in this study: Phonology §2.1.4, §2.1.3, §2.2.2.1; Boundary alternations §3.4.1 (glide insertion), §3.1.2.2 (glide fusion); Morphology §4.1.1.2, 4.1.5.2 (noun morphology), §4.2.4.1 (verb morphology).

⁹¹ The list reflects the general patterns observed; exceptions and dialectal variations may exist, e.g., Kambaata dialectal initial *w* ~ *y*. Consonant loss is sometimes compensated for by epenthetic sounds not listed. Segments and sequences like *c*, *j*, *p'*, *d'*, *ʔm*, *ʔn*, *ʔl*, and *ʔr* result from morphophonemic processes and contact.

PHEC	Burji	PNHEC	Had	Kam	Sid	Ged	Environment	Confidence
*k	k	*k	k	k	k	k	all	secure
*g	g	*g	g	g	g	g	all	secure
*k'	k'	*k'/ø	k'/k/ø	k'/k/g/ø	k'/k/h/ø	k'/h/ø	varies	arguable
*ʔ	ʔ/ø	*ʔ/ø	ʔ/ø	ʔ/ø	ʔ/ø	ʔ/ø	varies	secure
*h	h/ø	*h/ø	h/ø	h/ø	h/ø	h/ø	varies	secure
*ħ	h, y	*ħ	w, h	w/y, h	w, ø	w, ø/y	__*u, V__	arguable
*t'	t'/d'/c'	*t'	t'	t'	t'/d'/c'	t'/d'/c'	varies	reasonable
*c'	c'	*c'	c'	c'	c'	c'	all	reasonable

Reflexes of both ***l** and ***r** in Hadiyya are **l** in initial position and **r** in intervocalic position. The symbol **ø** indicates sound loss, as seen in **h/ø**, **w/ø**, and **y/ø**, which are examples of the loss of the glides in certain contexts, especially in non-initial position. The cases of **b/β/w**, **ʃ/s**, and **k'/k/g/ø** are examples of variation in reflexes depending on various contexts.

The geminates **bb**, **ll**, **rr**, **kk**, and **zz** are found in limited attestation as examples of regular (geminate) consonant sequences. Most, if not all, the clusters result from morpheme boundary alternation formations not part of the root. The same applies to a number of geminates.

The irregular **d'** occurs in Burji, Gedeo, and sometimes Sidaama, with correspondence to **t'** in Hadiyya and Kambaata in certain cognate sets. Many instances of **d'** in Burji, Gedeo, and Sidaama have direct parallels in Oromo. Sidaama shows free variation between **d'** and **t'**, correlating to Oromo influence. Non-initial **d'** often represents a reflexive suffix, which is also found in Oromo.

	Burji	Had	Kam	Sid	Ged	Environment
*t' →	d'	t'	t'	d'/t'	d'	#_
*t' →	c'	t'	t'	c'/t'	c'/t'	varies

The other irregular reflexes of ***t'** appear as **c'**, mainly in Burji and sometimes in Sidaama and Gedeo, with **t'** in Hadiyya and Kambaata. The distribution of these reflexes shows evidence of sibilant harmony in Burji (e.g., [18] *c'eeji* 'blood'); irregular correspondences with **c'** and **t'** in Sidaama and Gedeo; and parallels with areal sibilant patterns seen in **s** ~ **ʃ** variation (see Hayward, 1997, pp. 115–116).

Below is a list of selected common sound changes in HEC.

PHEC → reflex	Environment	Example	Confidence
*l → Had r	V_V	[137] *lal- → Had <i>laro</i> 'cattle'	secure
*r → Had l	#_	[184] *reñ ʔ → Had <i>leh-</i> 'to die'	secure
*rr → Had ll	_[non-initial]	[110] *barr- → Had <i>balla</i> 'day'	secure
*s → NHEC j	┘	[406] *wees-f- → Kam <i>weeffu</i> 'ensete tree (Ensete ventricosum)'	secure
*z → Kam z, rest of HEC d	all	[118] *zaak- → Kam <i>zaaz-</i> 'flow', Ged <i>daak-</i> 'swim'	secure
*t' → Burji, Ged, Sid d'	sporadic	[250] *t'iib- → Burji, Ged <i>d'iib-</i> 'push'	tentative
*b → β/w	V_V	[3] *aabb- 'father' → Sid <i>aaβa</i> 'father'	reasonable
*b → ø	V_V	[186] PHEC *c'ub- 'dip in, immerse' → Sid <i>c'uuʔ-</i> 'baptise'	secure
*h → ø	V_	[38] PHEC *liñ- → Burji <i>lia</i> , Sid <i>lee</i> 'six'	secure

Contact influence and cross-syllable obstruent harmony (fn. 36) explain some irregular changes.

Process	Scope	Diagnostics	Example	Confidence
Oromo *s → f	Sporadic, contact- induced	Irregular correspondences	[20] *sazeñ- → Burji <i>fadiya</i> 'three'	secure
Oromo verbal suffix -at	Verbal morphology	Correspondence with Oro -at forms	[450] Sid <i>saalfat-</i> 'be shy'	secure
d'/t' variation	Lexical items	Free variation in the same language	[417] Sid <i>d'ibba</i> ~ <i>t'ibba</i> 'disease'	secure
Sibilant harmony	C ₁ ...C ₂	t' → c', z → s where C ₁ is a voiceless obstruent, C ₂ is a sibilant	[18] *t'iig- → Burji <i>c'eeji</i> 'blood'; [20] *sazeñ- →	reasonable

Process	Scope	Diagnostics	Example	Confidence
			Kam 'three'	<i>sas-</i>
Velar harmony	C ₁ ...C ₂	t' → k' where C ₁ is a voiceless obstruent, C ₂ is a velar.	[18] → Kam 'blood', §2.1.2	<i>*t'iig-</i> <i>k'eg-</i> reasonable

Burji differs from the NHEC languages in several key aspects: (a) unlike NHEC languages, Burji shows evidence of a fossilized **d'**- prefix in verbs like *d'irk'*- (intr) vs. *rik'k'*- (tr) 'return'; (b) Burji consistently shows **d'** where NHEC languages show **t'** in certain cognate sets; (c) Burji shows evidence of Oromo-specific sound changes like ***s** → **f**. These distinctive features highlight the early separation of Burji from the NHEC languages and its constant contact with Oromo.

2.5.3 Vowel System

The Proto-Highland East Cushitic vowel system consisted of five vowels with a phonemic quantity contrast. The quantity contrast is analyzed moraically, with single vowels (V) carrying one mora and double vowels (VV) carrying two morae.

Table 5: *Proto-Highland East Cushitic Vowels*

	Front		Back
High	*i, *ii		*u, *uu
Mid	*e, *ee		*o, *oo
Low		*a, *aa	

This vowel system is remarkably stable across the HEC languages, with consistent reflexes in all languages except for some contextually conditioned variations (§2.2). The distinction between single and double vowels is phonemic, as evidenced by numerous minimal pairs in each language. The vowel ***e** is relatively infrequent compared to other vowels.

All HEC languages have terminal vowels (TVs), though their quality and quantity vary (§2.2.3). These TVs can be:

Double vowels (e.g., Sidaama *aroo* 'husband')

Single voiced vowels (e.g., Kambaata *saʔa* 'cow')

Single voiceless vowels (e.g., Burji *ʔiffi* 'she' (nom))

Terminal vowel quality often reflects case markers (nominative, accusative, genitive) or gender markers (§4.1.5). In HEC languages, terminal vowels have undergone various degrees of reduction through apocope, with Hadiyya, in which TVs are consistently voiceless (§§2.2.3, 2.4), showing the most advanced form.

Vowel doubling is a phonological process that often results from the loss of consonants, particularly glottals. For example, in Burji, the preconsonantal stem of *kaʔ-* is *kaa-* ‘get up’ through compensatory doubling after the loss of the glottal (§3.3.3). This process is observed in other HEC languages in specific alternations at morpheme boundaries.⁹²

2.5.4 Phonotactics and Prosody

PHEC syllable structure followed the pattern (C)V(V)(C)(C), with CV and CVC being the most common types. Initial consonant clusters are not permitted, while medial clusters are limited to two consonants and subject to sonority sequencing constraints. Some vowel-initial syllables in PNHEC derive from the loss of *kʼ* and *h* (§2.1.5). The HEC languages predominantly feature open syllables, with consonant clusters restricted to specific environments. Word-internal syllables may be closed by sonorant consonants or the first member of double consonants (§2.3.3). Consonant clusters are generally limited to geminate consonants, sonorant + obstruent sequences, and glottal stop + sonorant sequences (*ʔm*, *ʔn*, *ʔl*, *ʔr*). Metathesis patterns distinguish Burji (velar + sibilant → sibilant + velar) from NHEC (sonorant + glottal → glottal + sonorant), supporting the primary HEC subgrouping (§5.3.1). Non-initial position serves as an innovation environment for many consonants, allowing for greater diversity in consonant distribution. Therefore, the relevant phonotactic sections include §§2.3, 2.4 (Phonology), and the whole upcoming alternations at morpheme boundaries chapter (Ch. 3).

In Highland East Cushitic, accent/stress is associated with the most prominent syllable or a higher pitch. Accent placement in the HEC languages is primarily confined to the penultimate and ultimate syllables. While accent predominantly falls on the penultimate syllable synchronically, this is likely due to historical changes such as the reduction of double terminal vowels and

⁹² The HEC vowels are discussed in Phonology §2.2; Alternations at Morpheme Boundaries §3.4.1.1 (vowel epenthesis), §3.3.3 (compensatory doubling); and Morphology §4.1.5 (nominal suffixes).

apocope. Historically, the accent falls on the ultimate heavy syllable, which includes syllables with double vowels and closed syllables.

Accent placement interacts with vowel quantity and the phonetic status of the terminal vowel. Double vowels attract as accent more readily than single vowels, voiceless final vowels are unacceptable accent-bearing units, and in Hadiyya, the extensive reduction of final vowels effectively eliminates final accent in the material surveyed here. Taken together, these properties offer a plausible route from a historically final-heavy target to the synchronically widespread penultimate pattern.⁹³

The cross-referenced sections show that PHEC alternations at morpheme boundaries (Ch. 3) involves several systematic processes that affect the realization of phonemes in morphological contexts: common assimilation processes, especially involving sibilants and nasals, palatalization before front vowels in Burji and the glide **y* in the other HEC languages, affecting particularly alveolar consonants, epenthesis of vowels to break up impermissible consonant clusters, and epenthesis of nasals to compensate for degemination. The key morphological elements (Ch. 4) that trigger such morpheme boundary alternations include: the causative **-s*, which triggers assimilation and palatalization; the reflexive **-ʔ*, which interacts with stem-final consonants; and the singulative **-f*, which triggers palatalization.

The shared phonological innovations are documented in this chapter 2 and expanded in the following chapter 3. They constitute key isoglosses that, together with the morphological and lexical evidence presented in Chapter 4, establish Proto-Northern Highland East Cushitic as a genetic subgroup within HEC. The phonological evidence thus provides the foundation for the subgrouping proposed in §5.3.

⁹³ Details and discussion of HEC accent are found in Phonology §§2.3-2.4; Boundary alternations §3.2.2 (prosodic effects); and Morphology §4.1.4.4 (suprafix).