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

Sosal Altayeb Mohammed Ali, A.

4 Morphology

This chapter 4 reconstructs the nominal and verbal morphology of Highland East Cushitic and evaluates the distribution of morphological markers across the five HEC languages. The analysis follows the comparative framework established in Chapter 1 (§1.5) and is integrated with the phonological correspondences and reconstructions developed in Chapter 2. Each proposed marker is assessed in terms of its attestation, its allomorphy, and the environments by which it is conditioned, complemented by Chapter 3, along with its reconstructability to PHEC or to the PNHEC node.

Morphology is central to the interpretation of the boundary alternations documented in Chapter 3 because many processes are triggered by specific suffixes and stem classes. By making the relevant morphemes explicit and reconstructing them, this chapter provides a tighter account of why particular sandhi patterns recur across languages and why others are restricted to individual branches in HEC. The chapter also contributes evidence to the internal classification developed in Chapter 5 by identifying shared morphological innovations and by distinguishing inherited markers from branch-specific developments.

This chapter is organized as follows: Section 4.1 treats noun morphology, including number marking, nominal extensions, genitive constructions, and gender and case markers, with a separate subsection on personal pronouns. Section 4.2 turns to verb morphology and reconstructs the major derivational markers and support verbs. Section 4.3 synthesizes the reconstructed morphemes in a comparative summary and highlights the markers that are reconstructable.

4.1 Noun Morphology

This section surveys the main categories of noun morphology in HEC and reconstructs the markers that can be projected to PHEC or PNHEC. The discussion begins with number marking (§4.1.1) and nominal extensions (§§4.1.2, 4.1.3), then turns to genitive constructions (§4.1.4) and to gender and case morphology (§4.1.5). The final subsection addresses personal pronouns (§4.1.6).

Throughout, the aim is to separate broadly inherited patterns from branch-specific developments. For each marker, the analysis identifies its formal exponent(s), the environments that condition its allomorphy, and its

distribution across the five languages. Where nominal morphology triggers alternations at morpheme boundaries, cross-references to Chapter 3 are provided to contextualize the morphological description within the detailed discussion of the boundary processes.

4.1.1 Number

Highland East Cushitic number morphology distinguishes formally derived categories from functional ones. I use the terms singulative and plurative (Dimmendaal, 1983, p. 223; Corbett, 2000, p. 17) to refer to derived forms carrying overt number suffixes, whether marked against a base or against each other. Nouns may thus be formally unmarked, singulative, or plurative (see §1.5.2; Eyob Kelemework, 2015, p. 103), while singular and plural are reserved for function reference (Treis, 2008, p. 131).

4.1.1.1 Singulative

Alongside the suffix ***-ʃ** (realized as **-ʃo** with a terminal vowel), Hudson (1976, p. 251; 2006, p. 296) identifies an HEC singulative **-cco**, appearing as **-co** after sonorants and **-cco** after obstruents. Comparative evidence supports reconstructing these not as distinct suffixes, but as a single conditioned set. The proposed PHEC singulative allomorphs and their reflexes are summarized below.[148]

	PHEC	PHEC allomorphs	HEC language allomorphs
PHEC	***-ʃ →**	***-ʃ, *ʃ-ʃ**, ***-i-ʃʃ, *-c**	
Had	……		**-i-cco, -co**
Kam	……		**ʃʃu/ccu, -i-ccu**, **-cu**, **-ee-ccu**
Sid	……		**-ʃʃo**, **-i-cco**, **-co**
Ged	……		**-ʃo**, **-i-cco**, **-co**, **-jo**
Burji	……		**-ʃo, -ʃaa**, **-i-ccoo, -coo**

The distribution strongly supports a single singulative morpheme with phonologically conditioned allomorphs rather than multiple unrelated suffixes. The allomorphy is predictable: ***ʃ-ʃ** occurs after fricatives (via fusion); ***-i-ʃʃ** arises with epenthetic **-i-** after geminates or clusters; ***-ʃ** marks stem-final sonorants; and ***-c** follows stops. As argued in §3.1.1.5, the affricated allomorph ***-c**, originally restricted by context, appears to have expanded in the modern HEC languages at the expense of the post-alveolar allomorphs.

[148] Data sources and discussion: Burji (Tesfaye Baye, 2015, p. 95), Sidaama (Anbessa Teferra, 2000, p. 27), Hadiyya (Tadesse Sibamo, 2015, p. 48), Kambaata (Treis, 2008, p. 132), and Gedeo (Eyob Kelemework, 2015, p. 103).

In Gedeo, the additional singulative form **-jo** is treated here as a restricted voiced realization of the affricated allomorph (cf. §3.1.1.5 (Affricated **j**)), not as an independent singulative suffix.

A representative singulative example is PNHEC [190] **arr-* 'sun'. Sidaama *arriʃʃo* and Gedeo *ariʃʃo* 'sun' retain the post-alveolar allomorph ***-i-ʃʃ** in the environment of a geminate, whereas Hadiyya shows an affricated realization in *ellinco* 'sun'. Singulative marking is infrequently used in Burji (Hudson, 1976, p. 251).

4.1.1.2 Plurative

Plurative marking in HEC is expressed by several suffixes and one reduplication-based pattern. The suffix ***-uw-a** is found in Gedeo, Sidaama, and Hadiyya; ***-ey-e** is found in Gedeo, Sidaama, and Burji; and ***C_i~C_i-a** is found in Gedeo, Sidaama, and Kambaata. Burji and Kambaata instead commonly use the plurative suffixes **-na** and **-aakkata**, respectively.

Reduplication *C_i~C_i-a

Reduplication of the stem-final single consonant and addition of **-a** as a plurative formation is found in Gedeo (Eyob Kelemework, 2015, p. 107), Sidaama (Hudson, 1976, p. 252), and Kambaata (Treis, 2008, p. 135).[149] Hudson describes it as the most common form of plurative marking in Sidaama.[150] The following are examples of the reduplication-based plurative (***C_i~C_i-a**) alongside the corresponding unmarked (or singulative) base forms:

		Plurative	**Unmarked/sgv**
PHEC	***C_i~C_i-a**		
Gedeo	**n + a → n~na**	*man~n-a* 'men'	*man-jo* 'man / person'
	b + a → b~ba	*k'ub~b-a* 'fingers'	*k'ub-icco* 'finger'
Sid	**n + a → n~na**	*man~n-a* 'men'	*man-co* 'man'
	b + a → b~ba	*hogob~b-a* 'loads'	*hogob-o* 'load'
Kam	**n + a → n~na**	*min~n-a-ta* 'houses'	*min-i* 'house'
	s + a → s~sa	*boos~s-a-ta* 'water-carrying pots'	*boos-u* 'water-carrying pot'

[149] The **-a-** is described as a case marker in Kambaata followed by the feminine gender suffix **-ta** (Treis, 2008, p. 135).
[150] An alternative plurative suffix added to the same Sidaama stems is **-e** without geminating the non-geminate stem final or added to stems ending in a geminate.

Comparable reduplication-based pluratives are also attested elsewhere in Cushitic (see Zaborski, 1986; Ongaye Oda, 2013; Hayward, 1981; Wondwosen Tesfaye, 2020; Green, 2021; Tosco, 1991; Hetzron, 1976). On that basis, ***C_i~C_i-a** is plausibly a retention from PHEC; within HEC, however, the direct evidence currently comes from only three languages.

Plurative *-uw-a

The cognates of the plurative suffix ***-uw-a** are found in three HEC languages: Sidaama **-uwa** (Hudson, 1976, p. 252; Indrias Xa'miso et al., 2007, p. 635), Hadiyya **-uww-a** (Tadesse Sibamo, 2015, pp. 50, 84, 295), and Gedeo **-uwwa** (Eyob Kelemework, 2015, p. 102). The following examples illustrate the plurative suffix ***-uw-a**, showing the plurative form alongside the corresponding unmarked (or singulative) base:

		Plurative	**Unmarked/sgv**
PHEC	***C_iC_i-uw-a**		
Ged	**rr** + **uw-a** → **rr-uww-a**	*dubarr-uwwa* 'girl'	*dubarra* 'girl'
	n + **uw-a** → **n~n-uww-a**	*man~n-uwwa* 'persons'	*man-jo* 'person'
	nn + **uw-a** → **nn-uww-a**	*ann-uwwa* 'fathers'	*anna* 'father'
Sid	**nn** + **uw-a** → **nn-uw-a**	*ʔann-uwa* 'fathers'	*ʔanna* 'father'
	k'k' + **uw-a** → **k'k'-uw-a**	*sik'k'-uwa*, *sik'k'-ubba* 'stick' (pl)	*sik'k'o* 'stick'
Had	**l** + **uw-a** → **l~l-uww-a**	*k'ool~l-úwwa* 'clevers'	*k'oóra* 'clever'
	b + **uw-a** → **b~b-uww-a**	*ʔallab~b-úww-a* 'tongues'	*ʔallábo* 'tongue'

Across the examples, ***-uw-a** frequently co-occurs with stem-final consonant reduplication or glide gemination. Otherwise, it is attached to stems ending in a geminate, a cluster, or **ʔ**/**h**, e.g., Hadiyya *afoʔo*, *afoʔ-uwwa* (plv) 'hole' (Ritter, 2007, pp. 3–4). The next paragraph summarizes these distributional tendencies and exceptions.

Hudson (1976, p. 252) reconstructs ***-uba** (rather than ***-uwa**) largely on the basis of Sidaama, where intervocalic **b**, **β**, and **w** are allophonic (see §2.1.3 (PHEC *b)). He argues that stems with **b** encouraged the retention of the **b** in the suffix **-uba** (e.g., *buuda* 'horn', *buudubba* (pl)), which otherwise becomes **-uwa**. However, the distribution of Sidaama **-ubba**/**-uwa** is not limited to stems containing a labial **b**: for example, *sik'k'-uwa* and *sik'k'-ubba* 'stick' (plv) co-occur with stems lacking **b**, similarly, *ille* → *illubbà* and *illuwà* (plv) 'eye'. This weakens a strictly labial-conditioning account and leaves ***-uwa** as the more

economical comparative base and the variant with **b** as part of the allophonic alternations.

Intervocalic glide gemination may reflect a general strategy to stabilize a weak intervocalic glide, either via epenthesis and subsequent fusion (e.g., ***w** → ***w-y-** → **ww** /**V_V**) or via direct gemination. A parallel pathway is conceivable for ***-ey-e** (***y** → ***y-y-** → **yy** /**V_V**) on the assumption that the relevant glide behaves similarly in intervocalic position (see §2.1.4 (PHEC *w)). In Hadiyya and Gedeo, gemination plausibly contributes to the preservation of these sequences; in Sidaama, the intervocalic **b ~ β ~ w** alternation may instead allow the glide outcome without gemination.

Plurative ***-uw-a** is supported by cognates in Sidaama, Gedeo, and Hadiyya. It has no parallels in Kambaata or Burji. However, the plurative/collective ***-ey-e** is found in Burji, Sidaama, and Gedeo with no parallels in Kambaata (or Hadiyya).

Collective *-ey-e (f)

The cognates of the collective/plurative ***-ey-e** are found on adjectives and some nouns in Gedeo (Eyob Kelemework, 2015, p. 223), Sidaama (Kramer & Anbessa Teferra, 2020, p. 11), and Burji (Tesfaye Baye, 2015, p. 93). The following are examples of ***-ey-e** in Gedeo, Sidaama, and Burji, with the plurative/collective form next to the corresponding unmarked (or differently marked) base:[151]

		Plurative	**Unmarked/sgv**
Ged	**-eeye**	*c'im-eeye* 'old, experienced people'	*c'ima* 'old, experienced'
		korb-eeyye 'goats'	*korb-eessa* 'he-goat, billy'
Sid	**-e(e)yye**	*c'im-eeyye* 'arbitrators'	*c'im-eessa* (m), *c'im-eette* (f) 'arbitrator'
		hiyy-eyye 'wolf-like animal'	*hiyy-eyy-icco* 'wolf-like animal'
Burji	**-aayyee**/ **-aayyaa**	*c'im-aayyeé* 'elders'	*c'im-eyʃi̥* (m), *c'im-eyteé* (f) 'elder'
		ʔarʔ-eyyaá 'oxen'	*ʔarʔ-éyi̥* 'ox'

[151] Tesfaye Baye (2015, p. 93) glosses the Burji suffix as **-yyaa**, **-yyee**, but the comparative evidence suggests **-aayyaa**, **-aayyee** (cf. that *ʔarʔ-eyyaá* 'oxen' seems to be a concatenation of two suffixes given the existence of **ey** in the unmarked *ʔarʔ-éyi̥* 'ox').

The attested reflexes suggest that ***-ey-e** participates in gender/collective morphology. In dictionary sources, reflexes of ***-ey-e** are commonly described as feminine and/or collective in Gedeo (Wedekind, 2008), Sidaama (Gaspirini, 1983), and Burji (Roba Dame & Wedekind, 2008). Unlike plurative **-uw(w)-a**, these reflexes do not appear to be restricted to a particular suffixation environment.

The collective suffix ***-ey-e** is not found in Hadiyya or Kambaata. However, in Hadiyya, "[w]ith the exception of *tómmo* 'ten', *sáde* 'thirty', and *sóre* 'forty', the building morpheme for other morphologically complex tens (twenty, fifty, sixty, seventy, eighty and ninety) is [-*yye*] is (Lit:-ty)", e.g., *lamíyye* 'twenty', *ʔontáyye* 'fifty', *loháyye* 'sixty' (Tadesse Sibamo, 2015, p. 151; see comparative data in Treis, 2008, p. 305).

Both ***-uw-a** and ***-ey-e** are candidates for reconstruction to PHEC insofar as they have comparanda elsewhere in Cushitic (Zaborski, 1986; Ongaye Oda, 2013; Wondwosen Tesfaye, 2020; Mous, 1993; Mous, 2016; Harvey, 2018).

Miscellaneous Pluratives *-n, -r, -aakkata

Plurative *-n

The common plurative **-na** Burji (Tesfaye Baye, 2015, p. 90) and **-aano** in Sidaama (Anbessa Teferra, 2000, p. 39) may continue an older plurative ***-n**. This possibility is supported by (i) fossilized ***-n** forms elsewhere in HEC and (ii) the regular presence of **-n** in plural pronoun morphology.

The following list includes candidate fossilized reflexes of an older plurative ***-n** and the potential source stem without the plurative:

[142]	Fossilized plv *-n	Gloss	[254]	Source stem	Gloss
PHEC	***tam-n-**	'ten'	←	**-**	-
Burji	**tanna**	'ten'	←	***tun~tumm-** ← ***tum-**	'fist'
PNHEC	***tom-n-**	'ten'	←	**tontomma**	'fist'
Had	**tommo**	'ten'	←	**tuntummo**	'fist'
Kam	**-**	-	**-**	**tuntummu-ta**	'fist'
Sid	**tonne**	'ten'	←	**tuntummo**	'fist'
Ged	**tomme**	'ten'	←	**(kottoomoo)**	'fist'

PHEC [142] **tam-n-* 'ten' may reflect an older ***-n** formation related to the stem [254] **tun~tumm-* 'fist'. One possibility is that **tam-n-* represents a non-

reduplicated variant of the same stem complex; another is that [363] **omt-* ‘five’ is not derived directly from **tum-* ‘fist’, but via an older **tom~t-*, in which case **tam-n-* could reflect a plurative derivation within this set. The internal directionality remains uncertain.

Plurative -r

The trill **-r** is found as a plurative marker in three HEC languages: Gedeo has **-ra** (Wedekind, 2008, p. 2) or **-aro** (Eyob Kelemework, 2015, p. 223); Sidaama has in the genitive and relative formations a collective suffix **-r** (relative), **-re** (gen.acc), and **-ri** (gen.nom) (Anbessa Teferra, 2014, p. 42); and Kambaata has a nominalizing plural enclitic internally reconstructed as ***ra** (Treis, 2008, p. 362). In the rest of Cushitic, Kießling (1994, p. 53) has reconstructed a PWR collective suffix ***-arú**, e.g.,**maʕa-arú* ‘small red beans’.

Plurative -k

The plurative **-aakk-a-ta**, e.g., *yann-a*, *yann-aakk-a-ta* ‘proverbs’ (plv) is unique within Kambaata and its related languages. This plurative formation consists of the morpheme **-aakk**, the case vowel **-a**, and the feminine gender suffix **-ta** (Treis, 2008, p. 136). The ***-k-** and ***-t-** components of this Kambaata plurative form might have parallels in one EC language. Boni has a plurative suffix **-tə** and a preceding fused or retained **k**, e.g., Boni *míŋ*, *mín* ‘house’ has pluratives in Boni Jara *múŋ*, *múŋ-tə* (plv) and Boni Kilii *míŋ*, *míg*, *mig-oo-tə* (plv) ‘houses’. The final ***n** velarization to **ŋ** is due to assimilation to or fusion with a following **k**, which replaced the final **n** and became **g** in some examples, hence, a plurative **min-k-ta*. However, the **k** could also be posited for the unmarked variant *míŋ*. Boni Kilii *mig-oo-tə* (plv) ‘houses’ is comparable with the Sidaama plurative suffix *-oota* as in *man-co* (sgv), *mann-oota* (plv) ‘man’.

4.1.2 Deverbative

In addition to its singulative function, the suffix complex ***-ʃ** (with an affricated allomorph ***-c**) also derives deverbal nominals across HEC. As in §4.1.1.1, the comparative evidence supports a shared PHEC inventory (***-ʃ, *ʃ-ʃ**, ***-i-ʃʃ, *-c**), but in deverbatives the choice among these surface shapes is often conditioned by preceding concatenated suffixes. The following is a summary

of the reflexes of the PHEC deverbative allomorph set across the HEC languages[152]

	PHEC	**PHEC allomorphs**	**HEC languages allomorphs**
PHEC	***-ʃ →**	***-ʃ, *ʃ-ʃ, *-i-ʃʃ, *-c**	
Had	……		**-anca**, **-ca**, **-i-ʃʃa**
Kam	……		**-c-a(-ta)**, **-iʃʃ-a(-ta)**
Sid	……		**-aan-co, -an-ʃo**
Ged	……		**-an-jo**, **-acco**
Burji	……		**-aancoo**, **-issoo**

Taken together, the reflexes motivate a single deverbative morpheme with multiple phonologically and morphologically conditioned realizations, rather than independent deverbative suffixes. The Gedeo form **an-jo** is compatible with the same voicing pattern discussed for the singulative in §3.1.1.5 (Affricated **j**), i.e., a voiced realization of the affricated allomorph in a restricted morphological/phonological environment, rather than a distinct deverbative suffix. Gedeo has no allomorph of post-alveolar ***ʃ**, and Burji has an assimilated allomorph **ss**. Unlike the singulative, many allomorphs of the deverbative are conditioned by a preceding concatenated suffix. Hence, the allomorph ***-iʃʃ** is mainly found on stems with a suffix **-is**, so it is explainable as **-is+ʃ/c → -iʃʃ**. Although Burji can change **ʃ** into **s** unconditionally, a concatenated **-is** likely to have assimilated the deverbative **ʃ/c** fully into **ss** in **-issoó**. The suffixation conditions of the remaining PHEC allomorphs remain the same as those of the singulative.

The deverbative functions as a word-class changing suffix to derive nominals from (verbalized) stems (§4.2.5). The following are some examples:[153]

[152] Data sources and discussion: Kambaata (Treis, 2008, pp. 158–160), Sidaama (unsegmented in the source, which is Kawachi, 2007, pp. 317–318), Gedeo (Eyob Kelemework, 2015, p. 128), Burji (Tesfaye Baye, 2015, pp. 131–132), and Hadiyya (own interpretation based on data from Ritter (2007) and Tadesse Sibamo (2015, pp. 40–41)).

[153] Data sources and further discussion: Sidaama (Yri & Pepper, 2019; Kawachi, 2007, pp. 314–317; Indrias Xa'miso et al., 2007, p. 287), Gedeo (Eyob Kelemework, 2015, p. 128), Kambaata (Treis, 2008, pp. 158–160), Burji (Tesfaye Baye, 2015, pp. 127–132; Getahun Addo, 2017, p. 593), Hadiyya (Ritter, 2007).

	Verb	Gloss	Derived Noun	Gloss
Kam	*k'al-am-*	'be born'	*k'al-an-c-a*	'generation, birth'
	kicce?-	'be(come) sad'	*kiccecc-a*	'sadness'
	hut'-is-	'make shiver'	*hut'-iʃʃ-a*	'malaria'
Had	*k'ar-am-*	'be born'	*k'ar-an-ca*	'being born'
	kicce?-	'be sad'	*kiccecca*	'sadness'
	hut'i-is-	'cause to shake'	*hut'i-iʃʃa*	'malaria'
Sid	*had'-am-*	'to walk about'	*had'-án-co*	'wanderer'
	gan-am-	'to hit, beat, fight'	*gan-an-ʃo*	'fighting'
Ged	*d'ib-em-*	'be sick'	*d'ib-em-an-jo*	'patient'
	moor-	'to steal'	*mood'-acco*	'thief'[154]
Burji	*k'un-?-*	'pound'	*k'un-aancoo*	'grinding stone'
	gan?-	'grow'	*gann-issoo*	'growth'

Further examples with the post-alveolar **-ʃ** reflexes are discussed in §3.1.1.5 and §3.1.2.1.

The allomorph **-anc** is attested frequently across HEC. As Treis (2008, p. 160) accurately assumes, the frequent collocation of **-am** and **-c** likely constitutes a complex derivational device. Based on the semantics of a number of the derived nouns, they seem to be derived from a verbalized stem (§4.2.5), namely with suffixes as **-am**, **-ʔ**, **-is** + **-ʃ**.[155] Synchronically, Gedeo has a verbalized stem *t'uut'-em-* 'be sucked' from which the nominal *t'uut'-em-acco* 'something sucked' is derived (Eyob Kelemework, 2015, p. 132).

Yri & Pepper (2019) note that Sidaama words with such nominalizers may not have a stem without the derivational suffix **-co**, e.g., *fúr-co* 'door'. This is true for many derived lexemes. Therefore, Sidaama *fúr-co* 'door' was originally derived from the PHEC verb [143] **fur y* 'to open', even if the verbal stem is not attested synchronically (cf. Treis, 2008, p. 133).[156]

Finally, despite the formal resemblance, one should not expect the derivational deverbative suffix to fulfill a singular function besides the nominalizing function.

In addition to **-c** and **-ʃ**, three HEC languages have a deverbative with **-t**: Sidaama and Gedeo **-atto**, and Burji **-too** (feminine agentive/experiencer).

[154] The deverbative *mood'-acco* 'thief' is derived via an intermediate verb **moor-ʔ-*.
[155] Various deverbative markers are found attached to PHEC roots, such as [198] **tum-am-ʃ-*, **tum-t-iʃʃ-* 'blacksmith' from the verb [48] **tum-* 'to pound'. The reconstructed deverbative stems in this study do not all include the lexicalized denominative morphemes, but they should be considered upon further research.
[156] See Gedeo *farriyya* 'be clear, open' (Wedekind, 2008, p. 146).

The following examples from Gedeo and Burji illustrate this derivation, contrasting the base verb with the agentive or instrumental nominal:

	Verb	**Gloss**	**Derived Noun**	**Gloss**
Gedeo	*k'oc'-*	'to pick'	*k'oc'-atto*	'one who picks'
Burji	*k'un-*	'pound'	*k'un-too*	'pounder'
	ʔuum-	'create'	*ʔuun-too*	'creator'

While the derivational function is clear, the regular HEC correspondences for these specific variants require further investigation (see §3.1.1.5).

4.1.3 Problematic Extension

A small set of nominal stems shows irregular variation between forms with a glottal stop or fused glottalized consonants and parallel forms without glottalization. The conditioning for this glottalization is opaque, though the evidence permits several potential sources: (i) remnants of an archaic extension (possibly ***ʔ**), (ii) irregular or analogical glottalization, (iii) obstruent (glottalization) harmony (fn. 36), or (iv) multiple derivational pathways.

Comparative set [145] exemplifies this divergence, showing a split between Burji/Gedeo (glottalized) and the remaining HEC languages (plain).

[145]	**Glottal stop**	**No glottal stop**	Gloss
Burji	**ʃinʔaa**	**-**	'urine'
Had	**-**	**ʃuma**	'urine'
Kam	**-**	**ʃuma-ta**	'urine'
Sid	**-**	**ʃuma**	'urine'
Ged	**ʃiʔna**	**-**	'urine'

The glottal pattern in [145] suggests a morphological rather than purely phonological origin. Specifically, the glottalized variants likely reflect a cycle of cross-class derivation (noun → verb → noun). For instance, [145] may result from compounding with the support verb *ʔ* 'to do' (§§4.2.3, 4.2.2.1, 4.1.4.1), yielding a PHEC form like [170] **sum ʔ* 'urinate', which subsequently lexicalized as a noun in specific daughter languages.

A similar split appears in [163], where the glottal stop is retained in Burji and Hadiyya but absent in Kambaata and Sidaama, and a cross-class derivation with ***ʔ** 'to do' is not feasible.

[163]	**Glottal stop**	**No glottal stop**	**Gloss**
Burji	**(h)arʔeyi, aʔré**	**-**	‘ox’, ‘bull’
Had	**woʔla**	**-**	‘calf (of young cow)’
Kam	**-**	**worjaamu**	‘young bull’
Sid	**-**	**wor-aamo**	‘ox, bull’

Two further comparative sets, [46] and [180], illustrate the extent of this variation, highlighting both cross-language mismatches and language-internal doublets.

[46]	**Glottalized (C’)C’**	**Non-glottalized**	**Gloss**
Burji	**irk’a**	**-**	‘tooth’
Had	**ink’e**	**-**	‘tooth’
Kam	**ink’u-ta**	**-**	‘tooth’
Sid	**-**	**hinko**	‘tooth’

The complex glottal reflexes for this Common EC stem are discussed in §3.3.1.3 and §3.4.3.

[180]	**Glottalized (C’)C’**	**Non-glottalized**	**Gloss**
Burji	**-**	**utubaa**	‘pole, center-pole of house’
Had	**-**	**utuba**	‘pole, center-pole of house’
Kam	**-**	**utubu**	‘pole, center-pole of house’
Sid	**-**	**ʔutuba**	‘private property’
Ged	**utup’a**	**utuba**	‘support, e.g., forked stick’, ‘pole, center-pole of house’

Gedeo shows both glottalized *utup’a* and non-glottalized *utuba* reflexes for [180]. As both function as nominal stems, the glottalization likely stems from the derivational history, though Oromo *utubaa* ‘center-pole’ offers a potential contact source for the plain variant.

Two further sets illustrate comparable irregularities in glottalization.

> Had [81] *kobeʔe*, Sid *kop’é*, Ged *kop’e* ‘shoes’ (see Oro *kop’-ee*, *kob-ee* ‘sandal’);
>
> Kam [544] *c’up’i-ta*, Ged *c’uup’e* ‘dagger’ (cf. Amharic *c’ube* and Oro *c’uubee* ‘dagger’)

Finally, certain irregularities resolve as clear phonological repairs. The intervocalic glottal stop in Hadiyya [81] *kobeʔe* ‘shoes’ and [544] *c’ubeʔe* ‘dagger’, for example, functions as an epenthetic insertion to resolve vowel

hiatus (§3.3.4.1), distinguishing it from the opaque morphological glottalization discussed above.[157] The glottalized Kambaata and Gedeo **p'** in [544] is likely due to obstruent (glottalization) harmony (fn. 36), similar to the following example [49] (with full assimilation in Hadiyya):

> Had [49] *k'uunk'a*, Kam *k'uup'p'-iccu-ta*, Sid *k'uup'p'e*, Ged *k'uup'e*; *k'uup'p'e* 'egg(s)'.

Alternatively, the fused glottalized **p'p'** in [49] could be derived via the verbalizer **ʔ** (§4.2.5.1); for instance, Kambaata *k'uup'p'-iccu-ta* 'egg' might be derived from the nominal *k'ububb-eema* 'round, circular, ball-shaped' via the regular verbalizer **k'uup'p'-* 'be(come) circular' then used as a noun *k'uup'p'-iccu-ta* 'egg' (see the current verbalizer *k'ububb y-* 'be(come) circular)'.

In addition to the derivation explanation, the examples certainly include loan words; for instance, in [81], Sidaama *kop'é*, Gedeo *kop'e* with irregular single **p'** instead of the expected geminated glottalized **p'p'** (§4.1.3) might be a loan from Oromo, where both *kop'-ee* and *kob-ee* 'sandal' exist. The irregular single **p'** is also noticeable in [544] and [180]. However, the problematic examples cannot all be loans, for instance, [46] is attested throughout Cushitic usually as *ilk-* 'tooth'.[158]

4.1.4 Genitive

The genitive construction expresses possession or relational attributes between nouns in HEC. Across the five languages, genitive marking ranges from the conservative retention of a glottal consonant (Hadiyya) through vowel and suprafix systems (Sidaama, Kambaata) to morphologically complex clitic constructions (Burji, Gedeo). This section presents the comparative evidence and examines each genitive type: the glottal marker, vowel/suprafix systems, and the gender-clitic compounds (§4.1.4.1); gives an overview of the genitive pronouns (§4.1.4.2) and genitive components (§4.1.4.3); and concludes with a proposed historical scenario (§4.1.4.4). The reconstruction argues for a PHEC genitive suffix ***-ʔ**, whose reflexes are obscured by subsequent phonological processes but can be traced through Hadiyya glottal evidence,

[157] Similarly, **-ʔ-** epenthesis in Hadiyya [70] *afoʔo* 'hole' and Gedeo *afaʔo* 'mouth, language, opening', and **-y-** in Burji *afeyi* 'mouth, language' (§3.4.1.2).
[158] The Ts'amakko parallel is *ʔilǧe* 'tooth'.

Gedeo glottalized clitics, and compensatory doubling in Kambaata and Sidaama.

4.1.4.1 Genitive Markers

Highland East Cushitic languages use different forms of genitive morphemes depending on gender, case, and nominal type. The following is a summary of the principal genitive exponents attested on common nouns, proper nouns, and pronouns in each language. The **-**◌́ indicates a high-pitch suprafix without a segmental vowel, and **-C'C'** represents geminated glottalized consonants in Hadiyya.

	Common Nouns	**Proper Nouns**	**Pronouns**
Burji	**-ng**, **-nk** (m), **-nd**, **-nt** (f)	**-kk** (m), **-tt** (f)	**-kk** (m), **-tt** (f)
Ged	**-nk**, **-k** (m), **-nt'**, **-(t')t'** (f)	**-k** (m), **-t't'** (f)	**-k** (m), **-t't'** (f)
Sid	**-ú**, **-í** (m), **-te** (f), **-**◌́	**-i**, **-ø**	**-í** (m), **-é** (f)
Kam	**-V́(V)** (excluding **u**)	**-V́** (excluding **u**)	**-í** (m), **-é** (f)
Had	**-ʔ**, **-C'C'**, word order	**-**	**-t-ʔ** (m)

The markers show considerable formal diversity, and each is analyzed in the subsequent subsections. The reconstructive challenge is to identify the common source underlying this surface variation.

The ordering of the following subsections proceeds from what is hypothesized to be the most conservative trace of an ancient genitive marker (the glottal stop in Hadiyya), through intermediate stages involving phonological loss and change (vowel/suprafix systems in Kambaata/Sidaama), to the most morphologically complex and potentially innovative systems involving fused or layered clitics (Burji/Gedeo). This structure aims to build the case for the historical hypothesis presented at the end by showing the different outcomes across the language group.

Genitive *-ʔ

Tracing the history of HEC genitives, the reflexes involving a glottal stop, found predominantly in Hadiyya and also able to be deduced in Gedeo, potentially represent the remnants of an older genitive marking strategy, possibly involving a PHEC ***-ʔ** marker (cf. reflexive ***-ʔ** in §4.2.3.1). Examining

these forms provides a baseline for understanding subsequent innovations and diversifications across the language group.[159]

		Example	Gloss
Had	**-ʔ**	*maára* 'meat' + *mine* 'house' → *maaʔli míne*[160]	'butchery'
	-p'p'	*kitaaba* 'book' + *mine* 'house' *kitaap'p'i mine*	'bookstore'
	-k'k'	*elektiriiksiiteʔe* 'electricity' + *amat't'a* 'volume' → *elektriik'k'i ammat't'a*	'voltage'
	-t't'	*diinate* 'money' + *k'at'ooma* 'quantity' → *diinat't'i k'at'ooma*	'quantity of money'
Ged	**-t't'a** (f.acc)	*jissot't'a beltu*	'Jisso's girl'
	-t't'i (f.nom)	*manjiccotit't'i belto*	'the daughter of the man/woman'
	-nt'a (f.acc)	*warfint'a isso*	'dog's teeth'

Hudson (1976, p. 254) and Tadesse Sibamo (2015, p. 60) mention a glottalized (stem-)final sonorant when a noun is genitivized in Hadiyya. However, neither recognizes the glottal stop as a genitive extension or identify the source of such glottalization. The above is based on my interpretation of the available data from those sources and Ritter (2007), per comparable behavior of the glottal stop in such contexts in HEC (§3.3).

Thus, the glottal stop **ʔ** in the above examples appears in compounds and genitive phrases, even when absent from the base forms in Hadiyya. This suggests a morphophonological process, possibly triggered by a genitive marker **-ʔ**. In Gedeo, the glottalized **-t't'** (see the pronoun *iset't'e* 'her' (Eyob Kelemework, 2015, p. 267)) could also reflect a glottalized feminine marker **-t** resulting from fusion with a following glottal stop. Like Hadiyya, it might derive from a similar genitive **-ʔ**.

Treis (2008, p. 189) identifies a glottal element in the genitive and ablative forms as an unknown potential extension, e.g., *ber-é* (acc) 'yesterday', *ber-é ~ ber-e-ʔée* (gen), *ber-éecc ~ ber-e-ʔéecc* (ablative) (compare with the dative *ber-ée*) and *t'on-íta* (acc) 'before, recently', *t'on-é ~ t'on-e-ʔée* (gen). Later, she describes this glottal extension (at least for the genitive) as reminiscent of

[159] Data sources: Gedeo (Hudson, 1976, p. 254; Eyob Kelemework, 2015, pp. 101, 236); Hadiyya (Tadesse Sibamo, 2015, pp. 38, 60; Ritter, 2007, pp. 75, 82).
[160] Tadesse Sibamo (2015, p. 38) transcribes it as one word: *maaʔlimíne* 'butchery'. However, it is most likely a phrase, as consistently transcribed in the data from Ritter (2007), e.g., *laʔl mine* 'cattle area (in house)'.

the associative morpheme. A comparable glottal stop has been described as inserted before the derivational morpheme, e.g., *ceemm-í-ta* (acc), *ceemm-é* (gen) 'laziness', *ceemm-e?-aam-ú* (m.adj), *ceemm-e?-aam-í-ta* (f.adj) 'lazy' (Treis, 2008, p. 275). The latter analysis is the one adopted for all these glottal stops and has historical support in the form of epenthesis between three consonants (§3.3.4).

Genitive Vowel and Suprafix

The genitive systems in Sidaama and Kambaata use vowels and suprasegmental high pitch rather than the common overt consonantal markers. This likely reflects a distinct route of development characterized by the loss of an earlier consonantal element, potentially the ***-?** genitive or a subsequent clitic, followed by compensatory phonological adjustments such as vowel coalescence or accentuation, leading to the synchronic vowel and suprafix markers. The glottal stop is a common source of a high tone in various languages (Hombert et al., 1979, pp. 49–51).[161]

		Example	Gloss
Sid	**-u** (m Unmod)	*beett-ú faraʃʃ-i*	'The boy's horse'
	-i (m Mod)	*dangur-í beett-í faraʃʃ-i*	'Dangura's boy's horse'
	-i (m proper N)	*dangur-í faraʃʃ-i*	'Dangura's horse'
	-ó (f Mod)	*dangur-í beettó faraʃʃ-i*	'Dangura's girl's horse'
	-ó (f proper N)	*bulé faraʃʃ-i*	'Bule's horse'
Kam	**-á(a)** (f, m)	*roʃʃ-á (f) min-í*	'school'
		had-á (f) baad-ú	'abroad'
	-é(e) (f, m)	*Ludag-é (f) ball-úta*	'Ludage's mother-in-law'
	-ó(o) (f, m)	*az-ó (f) sa?-á*	'dairy cow'
	-í(i) (m)	*zaneecc-í (m) girk'-á*	'tusk of the / an elephant'

Sidaama employs vowels similar to those of the nominative, i.e., **-u** and **-i**, as a genitive marker of Unmodified (Unmod) and Modified (Mod) common masculine nouns, respectively.[162] The genitive **-i** is also attached to some

[161] Data sources: Sidaama (Kawachi, 2007, pp. 357–359), Kambaata (Treis, 2008, pp. 391, 383, 119). Some of these constructions are taken from longer contexts in the original sources, resulting in them potentially yielding different meanings or being difficult to translate when used in isolation. Therefore, the page numbers are provided for reference to the full context.

[162] For Kawachi (2007, pp. 247–248), "…"Modified" and "Unmodified" are capitalized because these terms are used here to refer to a morphosyntactic distinction specific to common nouns in Sidaama... Modified common nouns include not only those that have genitive NP dependents or adnominal dependents, but also those accompanied by the possessive pronominal suffix. Unmodified common nouns are those without any of

proper nouns, e.g., *dangur-í*. Additionally, the genitive vowels receive an accent as a genitive suprafix marker (Kawachi, 2007, pp. 356–357). The feminine common nouns, on the other hand, receive **-te** (Unmod), e.g., *beettó-te faraʃʃ-i* 'the girl's horse', and **-ø** (Modified and proper). When the feminine genitive suffix is **-ø**, the high pitch on the final vowel of the stem is analyzed as the sole genitive marker (Kawachi, 2007, p. 359) (e.g., *beettó*, *bulé* in the examples above).

Except **u**, all other vowels (**a**, **e**, **o**, and **i**) are attested in Kambaata marking the genitive, both as moraically single and double. In the genitive, they also receive a suprafix as an accent **-V́**, **-V́V** (Treis, 2008, p. 103). However, in contrast to Sidaama (Kawachi, 2007, p. 360), the accent on genitive nouns is sensitive to the length of a word in Kambaata; i.e., when an accent-neutral suffix extends the word, the accent shifts as far to the right as possible, e.g., the genitive *bit'-í* 'of a stick' → *bit'-i-ʔné* 'of your (plural) stick', *bit'-i-kkiʔné* 'of your (honorific) stick' (see also *gujj-í* → *gujj-i-sí* 'of his back', *ball-ó* → *ball-o-sé* 'of her mother-in-law', *hill-ée* → *hill-ee-ssá* 'of their intestine' (Treis, 2008, pp. 55, 107)).

Here is a summary of the similarities and differences between Sidaama and Kambaata vowel and suprafix genitives (**x** means attested, **-** indicates unattested):

	Short V	Double V	V(V) + Suprafix	Suprafix only	Suprafix shift
Sidaama	**x**	**-**	**x**	**x**	**-**
Kambaata	**x**	**x**	**x**	**-**	**x**

The two systems are too similar to be analyzed differently. The historical scenarios proposed below in §4.1.4.4 provide an alternative explanation that applies to both.

The vowels attested following the genitive clitics **=k-/=h-** and **=t-** below are usually associated with certain cases or predicative use.

these "Modifiers" (those neither with a dependent or with the possessive pronominal suffix)".

		=k-/=h-	**=t-**
Burji	Nominative	**-u**	**-i**
	Accusative	**-a**	**-a**
Ged	Nominative	**-i**	**-i**
	Accusative	**-a**	**-a**
	Predicative	**-e**	**-e**
Sid	Predicative clitic	**-o**	**-e**
	Clause final	**-**	**-i**
Kam	Clitic	**-a**	**-a**

The vowels following the **=k-** and **=t-** in Burji and Gedeo seem to be determined by the nominative-accusative distinction. Eyob Kelemework (2015, p. 338) notes that "[o]ne important aspect of Gedeo genitive case is that the form of genitive clitic shows variation in the nominative and accusative case. In the nominative case, the nominative case marking suffix -i would be suffixed, but -a is used in the accusative", e.g., the nominative *manjiccinki* becomes *manjiccinka* 'of the man' in the accusative. The same could be inferred from the Burji data. The accusative takes **-a**, but there are further vowel distinctions depending on the gender of the nominative: **-i** for the feminine and **-u** for the masculine.

The following sections distinguish between "Genitive Clitics **=(k)k**/**=(t)t**" and "Gender Clitics **=k**/**=t**". The former refers to the often phonologically fused or assimilated forms (e.g., geminates) integral to the main genitive constructions in languages like Burji and Gedeo. The latter term, "Gender Clitics", is used more broadly to discuss the **k**/**t** elements when they appear in other contexts or languages (e.g., Kambaata linkers/copulas, Sidaama predicative clitics) where their function as more independent agreement markers, linkers, or copula elements is clearer. This distinction helps highlight the potential functional origins and broader distribution of the **k**/**t** elements that became grammaticalized within the specific genitive structures of Burji and Gedeo.

Genitive Clitics =(k)k and =(t)t

Burji and Gedeo represent the most morphologically elaborate genitive systems in HEC. Both languages employ clitics **=k** (m) and **=t** (f), which interact with the genitive marker ***-ʔ** to yield geminated (**=kk**, **=tt**) or fused (**=t't'**) forms on proper nouns and pronouns, and epenthetic-nasal forms (**n=k**, **n=t**) on common nouns due to degemination. These processes suggest a stage involving the integration or fusion of the gender clitic with a preceding element, potentially the old genitive marker ***-ʔ**, as outlined in the final

hypothesis (§4.1.4.4). Sidaama retains a reflex only in the feminine **-te**. This paradigmatic complexity distinguishes the southern HEC languages from the simpler systems of Hadiyya and Kambaata.[163]

		Example	Gloss
Burji	**-kku** (m.nom)	*bariteékku ʔaábbu̯*	'Barite's father' (subj)
	-kka (m.acc)	*bariteékka̯ ʔaábboó*	'Barite's father' (obj)
	n=gu (m.nom)	*halamíngu̯ wóccu̯*	'(the) boy's dog' (subj)
	n=ga (m.acc)	*halamínga̯ woccoó*	'(the) boy's dog' (obj)
	-cci (f.nom)	*baritéccị̯ ʔaáyyị̯*	'Barite's mother' (subj)
	-tta (f.acc)	*baritéttа̯ ʔaáyyeé*	'Barite's mother' (obj)
	n=ci (f.nom)	*halamíncị̯ hárrị̯*	'(the) boy's donkey' (subj)
	n=da (f.acc)	*halamínda̯ harreé*	'(the) boy's donkey' (obj)
Ged	**=ki** (m.nom)	*jissoki belti*	'Jisso's son' (subj)
	n=ki (m.nom)	*manninki min-i*	'the people's houses'
	=ka (m.acc)	*belto jissoka*	'Jisso's son' (obj)
	n=ka (m.acc)	*ayyaaninka lukkicco*	'a hen for a holiday'
	=ke (m.pred)	*weesetike ficcani*	'weese-plant's farm'
	n=ke (m.pred)	*bisinke gogani*	'the skin of a body'
	=t'i (f.nom)	*jissot'i belti*	'Jisso's daughter' (subj)
	n=t'i (f.nom)	*lamaʔint'i beltanna*	'the young of a rat' (subj)
	=t't'a (f.acc)	*belto jissot't'a*	'Jisso's daughter' (obj)
	n=t'a (f.acc)	*hajyyomaint'a galca*	'position of power' (obj)
	=(t')t'e (f.pred)	*hak'iccit't'e lame dame-ʔnni*	'the two wings of a tree' (pred)
	n=t'e (f.pred)	*hak'iccint'e dame*	'tree's wing' (pred)
Sid	**-te** (f)	*beettó-te faraʃʃ-i*	'the girl's horse'

Burji and Gedeo morphemes with **-n-** are found on common nouns (§3.4.3) while the geminate variants are mostly found on proper nouns and pronouns. Burji morphemes **n=gu** (m.nom), **n=ci** (f.nom), **n=ga** (m.acc), and **n=da** (f.acc)[164] are found on common nouns, whereas **-kku** (m.nom), and **-cci** (f.nom) are found on proper nouns (see §3.2.1 for palatalization and §3.1.1.1 voicing assimilation).

For Gedeo, Hudson (1976, p. 254) has **-nka** and **-nt'a** (accusative), **-nki** and **-nt'i** (nominative), e.g., *lalinka goga* 'a cow-hide' and *warʃint'a isso* 'dog's

[163] Data sources: Burji (Tesfaye Baye, 2015), Gedeo (Eyob Kelemework, 2015), Sidaama (Kawachi, 2007).

[164] Tesfaye Baye (2015) segments the Burji morphemes as suffixes: **-ngu** (m.nom), **-nci** (f.nom), **-nga** (m.acc), and **-nda** (f.acc). The above clitic-based segmentation (with epenthetic **n**) is used for ease of reading, since this matches the pattern in the rest of the HEC languages.

teeth' (see section on Nasal **n** below). Eyob Kelemework (2015, p. 270) specifies the differences as follows: **=ki, =(t)t'i** (nominative), **=ka, =(t')t'a** (accusative), and **=ke, =(t')t'e** (predicative), with a difference in the vowel quality, and **k** on the masculine and **t** on the feminine forms. Gedeo also shows a unique switch (apposition) in the order of the possessor and possessed in the object genitive constructions as discussed in §4.1.4.3, e.g., *belto jissoka* 'Jisso's son' (obj) vs. *jissoki belti* 'Jisso's son' (subj) (see Eyob Kelemework, 2015, p. 236).

Sidaama **-te** is found on Unmodified feminine common nouns. The genitive of the Modified common nouns is marked with a high-pitch suprafix **-**◌́ on the stem-final vowel only.

Gender Clitics =k and =t

While the previous section examined the heavily integrated and often phonologically altered forms of the **k**/**t** clitics in Burji and Gedeo genitives, similar elements also appear in other constructions across HEC, notably in Sidaama and Kambaata, often functioning more clearly as independent gender/agreement markers, linkers, or even copulas. Analyzing these functions provides crucial context for understanding the potential origin and functional range of the **k/t** elements that became integral parts of the genitive constructions in languages like Gedeo and Burji.

Sidaama uses the clitics =**h** (m) and =**t** (f) as dependent morphemes added to genitive constructions. In Kambaata, "[t]he possessive suffix … is attached after the constituent consisting of head and -ha / -ta ... An obligatory linker morpheme -n is inserted between -ha / -ta and the possessive suffix ... Predicate noun, -ha / -ta, and possessive suffix form a single phonological word" (Treis, 2008, p. 418). Hudson (1976, p. 254) describes **-ha** (**-ka**) (m) / **-ta** (f) as gender or 'article' suffixes that occur in genitive constructions with pronoun possessors in Kambaata. Note that Treis (2008, p. 418) does not identify **-ha** (**-ka**) (m) / **-ta** (f) as articles but rather as a copula or a linker.[165]

[165] Anbessa Teferra (2014, pp. 34–36) claims that the suffixes **-k** (m) and **-t** (f) are found on pronouns and common/proper nouns (as) genitive constructions in Sidaama. However, such analysis is challenged when compared to Kawachi's analysis of Sidaama.

		Example	Gloss
Sid	**=ho** (m)	*kuni mini danguri ann-i=ho*	'this house is Dangura's father's'
	-te (f)	*beettó-te faraʃʃ-i*	'the girl's horse'
	=ti (m/f)	*hakku faraʃʃo-se=ti*	'that is her horse'
Kam	**=ha** (m)	*níi báad-i=a gaʃʃáancu*	'(it is) the ruler of our country'
	-ta (f)	*aaz-éen-ta-s*	'inside of it'

The genitives of Sidaama could be classified into attributive (adnominal), e.g., *beettó-te faraʃʃ-i* 'the girl's horse', and predicative. The predicative genitive constructions receive **=te** (f) and **=ho** (m), i.e., predicative noun-phrase clitics (npc.pred) in Kawachi's (2007, p. 250) words. The clitics are suffixed to the Modified common nouns, proper nouns, and pronouns, e.g.:

kuni mini danguri ann-i=ho
Dangura.gen.m father-gen.m=npc.m.pred
'This house is Dangura's father's.'

kuni mini isí=ho
3sg.m.gen=npc.m.pred
'This house is his.'

hatti ooso dangur-í=te
Dangura-gen.m=npc.f.pred
'Those children are Dangura's'

They are also suffixed to the Unmodified common nouns preceded by the suffix **-nni**, e.g.:

kuni mini t'agisaanc-u-nni=ho
healer-gen.m-nni=npc.m.pred
'This house is the healer's'

Additionally, **=ti** occurs clause-finally for both masculine and feminine, such as on the pronoun (Kawachi, 2007, pp. 250–251, 356, 247).

hakku faraʃʃo-se=ti
horse-3sg.f.poss=npc.pred
'that is her horse'

Gedeo equivalents of predicative noun-phrase clitics in Sidaama are **=te**/=**t't'e** (f.sg/pl), **=ke** (m.sg/pl), and **=ra** (pl), which are also labeled as NPCs (Eyob Kelemework, 2015, p. 298). Gedeo NPCs are indistinguishable from the

predicative genitives **=ke** and **=(t')t'e**, hence the proposed classification of genitives in Sidaama into attributive and predicative.

However, Anbessa Teferra (2014, pp. 34–36) analyzes the predicative genitive construction in Sidaama differently. The genitive suffixes of the proper nouns, common nouns, and pronouns in such constructions include, for the nominative, **-h-u** (gen.m-nom), **-t-i** (gen.f-nom), and **-r-i** (gen.pl-nom) in the following examples:

beetto-nni-h-u
girl-perp-gen.m-nom
'the girl's' (m.subj)

beetto-nni-t-i
girl-prep-gen.f-nom
'the girl's' (f.subj)

beetto-nni-r-i
girl-prep-gen.pl-nom
'the girl's (pl.subj)'

Additionally, they include, for the absolutive (or accusative), **-ha** (gen.m.abs), **-ta** (gen.f.abs), and **-re** (gen.pl.abs) in the following examples:

beetto-nni-ha
girl-prep-gen.m.abs
'the girl's (m.obj)'

beetto-nni-ta
girl-prep-gen.f.abs
'the girl's (f.obj)'

beetto-nni-re
girl-prep-gen.pl.abs)
'the girl's (pl.obj)'

The following is a summary of the differences in analyses of Sidaama:

Feature	Kawachi (2007) Analysis	Anbessa Teferra (2014) Analysis
Core Markers	Predicate noun-phrase clitics: **=ho** (m), **=te** (f)	Genitive suffixes: **-h-u** (m), **-t-i** (f), **-r-i** (pl)
Marker Final Vowel	Part of the clitic	Case markers (nominative/absolutive)
Stem-final Vowel	Genitive marker (vowel/suprafix)	Part of the stem
Conditioning	Based on noun modification status (Modified/Unmodified)	Based on case (nominative/absolutive)
-nni Element	Suffix preceding clitic on Unmodified common nouns	Preposition preceding genitive marker
Plural Marker	Not explicitly part of this genitive system analysis	Genitive suffix **-r-** identified

In addition to the variation in the morphological analysis (Modified/Unmodified vs. nominative/absolutive), Anbessa Teferra's interpretation differs from Kawachi's in the description of the consonants and vowels, with differences even in their quality and assigned function. Kawachi interprets **=t** and **=h** as predicate clitics and the vowel following the stem or the suprafix (see the section on Genitive Vowel and Suprafix) as the genitive markers. Meanwhile, for Anbessa Teferra, **-t** and **-h** are the actual genitive markers, and the vowels are nominative/absolutive case markers. Anbessa Teferra's additional **-r** has been reported as animate **=re** (pl.acc), =**ri** (pl.nom/gen) in Sidaama (Kawachi, 2007, p. 243), and nominalizer **=ra** in Kambaata (Treis, 2008, p. 240). Therefore, the **-h-**, **-t-**, and **-r-** in Anbessa Teferra's examples might be nominalizers, e.g., 'the girl's (one(s))' (Treis, p.c.).[166]

In Kambaata, **-ha**/**-ta** are found in more complex nominal constructions involving the genitive and possessive pronouns. Treis (2008, pp. 351–352, 418) analyses them as enclitic, linker (l), and copula (cop2) depending on the nominal construction, e.g.:

locative: aaz-éen-ta-s
interior-m.loc-l-3sg.m.poss
'inside of it' (p. 195)

[166] "Headless genitive nouns are combined with the nominalizer =ra ...."
maal-i=r-a
meat-m.gen=ra-m.acc
'the ones of meat' (here: the hens that are suitable for meat production) (Treis, 2008, p. 240)

dative: hiz-éeha-n-se
sibling-m.dat-l-3sg.f.poss
'to her brother' (p. 226)

oblique: bahír-ua-n-s hiz-éeha
elder-m.obl-3sg.m.poss sibling-m.dat
'to his elder brother' (p. 226)

instrumental comitative perlative (icp): ʃiin-íin-ka-s
side-m.icp-l-3sg.m.poss
'his side' (p. 209)

ísi-ta béet-u
3sg.m.gen-f.cop2 child-f.pred
'(it) is his daughter' (p. 418)

predicate: nίi báadi-a gaʃʃ-aan-c-ú
1pl.gen country-m.gen-m.cop2 rule-aan-sg-m.pred
'(he) is the ruler of our country' (p. 414)

mín-i-a-n-se-baʔa
house-m.pred-cop2-l-3sg.f.poss-neg
'(it) is not her house' (p. 352)

The examples show the lenition of **k** → **h** → **ø**, which is confirmed by Treis' (2008, p. 352, fn. 31) observation that "[t]he use of a (masculine) linker morpheme -ka after masculine nouns is not rejected as ungrammatical, but it is hardly attested in the corpus".

	masculine **-k-**	feminine **-t-**	plural **-r-**
Ged	**=ke** (npc)	**=te**/=**t'e** (npc)	**=ra** (npc)
Sid	**=ho** (npc)	**=te** (npc)	**=ri**
Kam	**-ha**	**-ta**	**=ra** (nominalizer)

In addition to the nominal functions, the terminological consensus in Gedeo (Eyob Kelemework, 2015, p. 270) and Sidaama (Kawachi, 2007, p. 250) is that **=k**/**=h** and **=t** are (en)clitics. Therefore, clitic is used here as a cover term wherever those exponents are used with the genitive. The clitic analysis is supported by the fact that, although the modifier receives the clitic, the gender of the specific clitic must agree with that of the head noun (see the subsection on Gender) at least in Gedeo and Burji. The Kambaata suffixes **-h**/**-t** are most

likely unrelated to the clitics of Gedeo and Burji and are instead similar to the common exponents found in other nominals (see §4.1.5.1), including the prominent feminine accusative **-ta**.

Offering a compatible historical perspective, Tosco (1994) proposes a specific origin for these widespread **k**/**t** elements, which he terms “Gender Markers”. He argues they derive from grammaticalized copulas used in originally left-dislocated cleft constructions (Modifier + Copula + Head). Over time, the copula, following the modifier, was reinterpreted as an agreement marker attached to the possessor but agreeing with the possessed noun. Tosco finds evidence for this copular origin in the formal similarities between these markers and overt copulas still present in several HEC languages. He further suggests these **k**/**t** elements might have initially functioned more broadly as determiners or invariant copulas before being specialized for gender marking.

Nasal n

A further layer of complexity, primarily affecting common nouns in Burji, Gedeo, and specific Sidaama constructions, involves the appearance of a nasal element **n** preceding the **k** or **t** clitic. This nasal is arguably not part of the original clitic structure but rather a later phonological insertion (epenthesis) that arose following the simplification (degemination) of earlier fused geminate forms like **kk** and **tt**, preserving a trace of intermediate geminated consonants (§3.4.3).

		Example	Gloss
Burji	**n=gu** (m.nom)	*halamíngụ wóccụ*	‘(the) boy’s dog’ (subj)
	n=ga (m.acc)	*halamínga woccoó*	‘(the) boy’s dog’ (obj)
	n=ci (f.nom)	*halamíncị hárrị*	‘(the) boy’s donkey’ (subj)
	n=da (f.acc)	*halamíndạ harreé*	‘(the) boy’s donkey’ (obj)
Ged	**n=ki** (m.nom)	*manninki min-i*	‘the people’s houses’
	n=ka (m.acc)	*ayyaaninka lukkicco*	‘a hen for a holiday’
	n=ke (m.pred)	*bisinke gogani*	‘the skin of a body’
	n=t’i (f.nom)	*lamaʔint’i beltanna*	‘the young of a rat’ (subj)
	n=t’a (f.acc)	*hajyyomaint’a galca*	‘position of power’
	n=t’e (f.pred)	*hak’iccint’e dame*	‘tree’s wing’
Sid	**nni=ho** (m)	*t’agisaancunniho*	‘(it is) the healer’s’
	nni=ti (f)	*woʃiccunniti wajjido lekka*	‘The dog’s (m) white (f) legs (f)’

Burji **-ngu** (m.nom), **-nci** (f.nom), **-nga** (m.acc), and **-nda** (f.acc) are found on common nouns, corresponding to **-kku** (m.nom), **-cci** (f.nom) on the proper nouns (see §3.2.1 for palatalization and §3.1.1.1 voicing assimilation). For

Burji, Hudson (1976, p. 254; 1981, p. 102) has voiceless **-nka**, **-nta** (which could represent a dialectal difference), and no vowel following **-kk** or **-tt** (but there are devoiced vowels in Tesfaye Baye (2015)). In Gedeo, instead of emphatic **-n** and **=t'e** / **=t'i** / **=ke** / **=ki** / **=ka** (Eyob Kelemework, 2015, p. 117), Hudson (1976, p. 254) has **-nka** and **-nt'a** (accusative), **-nki** and **-nt'i** (nominative) on the common nouns, e.g., *lalinka goga* 'a cow-hide' and *warfint'a isso* 'dog's teeth'.

The **n** before **k** and **t** is described as a linker by Hudson (1981, pp. 102, 108) in Burji and Gedeo and emphatic in Gedeo (Eyob Kelemework, 2015, pp. 116–117, 338).[167] In Sidaama, **-nni** is described as an obligatory definite suffix preceding the clitics of the Unmodified genitive noun (Kawachi, 2007, p. 384), while Anbessa Teferra (2014, p. 36) identifies it as a postposition marker.[168]

For Tosco (1994), however, the interpretation of this **n** element in Burji and Gedeo differs from the above synchronic linker or emphatic marker descriptions. He proposes that it represents a historically later determinative element, introduced specifically with common nouns after the original determinative function of the **k**/**t** markers themselves had faded. The typical absence of this **n** with inherently definite proper nouns and possibly the definite suffix **-nni** in Sidaama (Kawachi, 2007, pp. 77–78) supports this interpretation of **n** as a secondary definiteness marker. However, the analysis adopted here reinterprets this sequence: The geminates ***kk** and ***tt** result from the fusion of ***ʔ=k** and ***ʔ=t** (genitive + clitic), with **-n-** arising upon degemination to maintain the biconsonotal construction.

Overall, the genitive marker ***-ʔ** is reconstructed at the PHEC level as arguable: Direct evidence survives in Hadiyya (overt glottal stop and **C'C'** glottalization), Gedeo (glottalized **t't'** feminine), and indirectly in Kambaata and Sidaama (compensatory vowel doubling and accentuation following glottal loss). Burji and Gedeo additionally preserve the gendered clitics ***=k** (m) and ***=t** (f), which represent concatenated suffixes attached to the genitive construction. The proto-status of the combined construction ***-ʔ=k/t** is tentative, given that Hadiyya and Kambaata show no synchronic trace of the velar/dental clitics.

---

[167] According to Treis (2008, pp. 418, 220), Kambaata has a linker **n** before the possessive suffix **-s**, and the emphatic function of **n** is also discussed under the pragmatically determined **-n** morpheme.

[168] Anbessa Teferra's (2014, p. 35) use of the term postposition for **-nni** is vague in such a context. Is it an instrumental or ablative marker in such a construction?

4.1.4.2 Possessive / Genitive Pronouns

Personal possession is expressed by several constructions across the HEC languages. In Burji and Gedeo, third-person possessive pronouns show the same gender-conditioned endings that recur in nominal genitive constructions (§4.1.4.1), suggesting these forms are morphologically complex. Hadiyya similarly uses independent third-person possessive pronouns preceding the possessed noun. In Kambaata and Sidaama, by contrast, third-person possession is marked both by a pronoun and by a suffix on the possessed noun itself (**-si** for masculine, **-se** for feminine possessors).

The list below presents the third-person possessive/genitive forms with illustrative phrases showing their syntactic distribution.[169]

	Genitive	**Example**	**Gloss**
Burji	**ʔisákkụ** / **ʔisákkạ** (3sg.m.nom/acc)	*ʔisákkạ ʔaabboó*	‘his father’
	ʔiʃeékkụ / **ʔiʃeékkạ** (3sg.m.nom/acc)	*ʔiʃeékkạ ʔanaá*	‘her husband’
	ʔisáccị / **ʔisáttạ** (3sg.f.nom/acc)	*ʔisáttạ ʔámạ*	‘his wife’
	ʔiʃeéccị / **ʔiʃeéttạ** (3sg.f.nom/acc)	*ʔiʃeéttạ ʔayyé-ccị*	‘her mother’
Ged	**isike** / **isika** (3sg.m.pred/acc)	*isika belto*	‘his son’
	iseke / **iseka** (3sg.m.pred/acc)	*iseka belto*	‘her son’
	isit’t’e / **isit’t’a** (3sg.f.pred/acc)	*isit’t’e beltoni*	‘his daughter’
	iset’t’e / **iset’t’a** (3sg.f.pred/acc)	*belto iset’t’a*	‘her (own) daughter’
Had	**ʔít’t’i** (3sg.m)	*ʔít’t’i lókko*	‘his leg’
	ʔísi (3sg.f)	*ʔísi lókko*	‘her leg’
	ʔíssi (3sg.hononrific)	*ʔíssi lókko*	‘his/her leg’ (honorific)
Sid	**isí** (3sg.m)	*isí ama*	‘his mother’
	isé (3sg.f)	*isé beettó*	‘her child’
Kam	**isí** (3sg.m.gen)	*ísi-ta béetu*	‘it is his daughter’
	isú (3sg.m.acc)	*isí beetú*	‘his son’
	isé (3sg.f.gen)	*isé beetú*	‘her son’
	iséta (3sg.f.acc)	*isé beetú-ta*	‘her daughter’

[169] Data sources: Hadiyya (Tadesse Sibamo, 2015, p. 88), Burji (Tesfaye Baye, 2015), Gedeo (Eyob Kelemework, 2015. pp. 269, 115), Sidaama (Kawachi, 2007), and Kambaata (Treis, 2008, p.c.).

Gender agreement in possessive constructions varies by language: Some mark possessor gender (Sidaama and Kambaata), and others mark possessed noun gender (Burji and Gedeo) (see §4.1.4.3 for a detailed analysis). In Gedeo, these possessive forms combine with person- and case-conditioned enclitics. Eyob Kelemework (2015, p. 270) identifies three sets: **=ki, =(t')t'i** (nominative), **=ka, =(t')t'a** (accusative), and **=ke, =(t')t'e** (predicative), yielding the surface variants shown in the examples above. Kambaata and Sidaama instead mark possession on the possessed noun with the suffixes **-sí** (3sg.m) and **-sé** (3sg.f).

The first-person possessive form may reflect a compound of the genitive marker **ʔ** and a first-person pronominal element **-y**. This same **-y** element appears in other first-person pronouns (§4.1.6.1), supporting its analysis as a core person morpheme. Forms preceded by a hyphen are suffixal variants:

	[135]		[188]	
PHEC	***ʔ-y**	'my'	***ni-ʔ-**	'our'
Burji	**ʔíyyụ / ʔiíyyạ**	'my'	**ninka/ninku, ninci/ninta**	'our'
PNHEC	***ʔy**	'my'	***niʔ-**	'our'
Had	**ʔi, -i**	'my'	**ni**	'our'
Kam	**íi, -ʔe**	'my'	**nii, -nne**	'our'
Sid	**-ʔya, ʔane**	'my'	**ninké, -nke**	'our'
Ged	**anki / ant'i**	'my'	**noʔoki, noʔot't'i**	'our'

These reconstructions remain provisional due to substantial diversity across daughter languages. Second-person possessive pronouns show similar patterns, with the fused (§3.3.1) and metathesized (§3.3.2) genitive glottal stop as well.

	[253]	2sg.poss	[187]	2pl.poss
PHEC	***ki-ʔ-**	'your' (sg)	***kin-ʔ**	'your' (pl)
Burji	**ʃiyyú/ʃíyyạ, ʃíccị/ʃíttạ**	'your' (sg)	**ʃínkụ/ʃínkạ, ʃínci/ʃíntạ**	'your' (pl)
PNHEC	***ki-ʔ, *-ʔ-ki**	'your' (sg)	***kin-ʔ-** → ***kiʔn-**	'your' (pl)
Had	**ki-**	'your' (sg)	**kiʔnúwwi**	'your' (pl)
Kam	**kíi, -kki**	'your' (sg)	**kiʔneʔee, kiʔnee, -ʔne**	'your' (pl)
Sid	**-kki**	'your' (sg)	**kiʔné**	'your' (pl)
Ged	**ati-ki, ati-t'i**	'your' (sg)	**(haʔnnoke, haʔnnot't'e)**	'your' (pl)

The Gedeo parallel stem *haʔn-* is generalized from the second-person plural nominative stem [104] *haʔno* 'you' (pl) (see §4.1.6.1).

The third-person pronouns also show the traces of the genitive glottal stop in their interaction with the consonant of the clitic via the regular processes of fusion in Hadiyya and Gedeo and gemination in Burji (§3.3.1).

	[279]	Poss	[366]	Poss	[318]	Poss
PHEC	***isa-ʔ-**	'his'	***ise-ʔ-**	'her'	***isn-ʔ-**	'their'
Burji	**ʔisákkụ/ ʔisákkạ, ʔisáccị/ ʔisáttạ**	'his'	**ʔiʃeékkụ/ ʔiʃeékkạ, ʔiʃeéccị/ ʔiʃeéttạ**	'her'	**ʔisinákkụ/ ʔisinákkạ, ʔisináccị/ ʔisináttạ**	'their'
PNHEC	***isi-ʔ-**	'his'	***ise-ʔ-**	'her'	***isn-ʔ-**	'their'
Had	**issi, ʔíṭ'ṭ'i**	'his/her'	**issi, ʔísi**	'her/his, her'	**ʔissúwwi**	'their'
Kam	**isí**	'his'	**isé**	'her'	**issoʔóo**	'their'
Sid	**-si**	'his'	**-se**	'her'	**-nsa**	'their'
Ged	**isike/isit'ṭ'e**	'his'	**iseke/iset 'ṭ'e**	'her'	**insaʔnnek e/insaʔnn et'ṭ'e**	'their'

The geminated **ss** in the plural Hadiyya cognate *ʔissúwwi* 'their' is from the full assimilation of the stem-final **s** and the plurative **-n** as discussed in §3.1.1.2. This plural stem was subsequently generalized to the singular stems in the masculine and the feminine, but is still transcribed (or alternates) with the single **s** in some sources.

4.1.4.3 Genitive Components

The previous sections discussed the genitive markers. The following paragraphs briefly present word order and gender assignment in genitive constructions.

Word Order

The typical order of the words in genitive constructions in HEC is that the possessor precedes the possessed (Modifier-Head). The possessor receives the genitive marker/clitic. This Modifier-Head (Possessor-Possessed) structure contrasts with the Head-Modifier order common in related Lowland East Cushitic languages. Tosco (1994) argues strongly that this Modifier-Head order in HEC is not original but represents a significant historical shift from an earlier Head-Modifier stage, likely in archaic Proto-East Cushitic. He proposes that this typological restructuring was driven by the grammaticalization of specific constructions, such as the cleft-derived genitives discussed earlier,

which harmonized HEC's nominal syntax with its consistent Verb-Final sentence structure.

In Hadiyya, Hudson's (1976, p. 254) "apposition" and Tadesse Sibamo's (2015, p. 60) "juxtaposition" are used to describe the order within the genitive components and assign it as the only way to express genitive relationships regardless of gender (cf. the subsection on Genitive ʔ). However, since the Modifier-Head is the default genitive component order in addition to the markers in the rest of HEC, word order terms such as apposition and juxtaposition do not tell us much comparatively.

gareec maara	'sheep meat'
ii-ann mentic	'my father's wife'
jallor ooso	'Jalloro's children'
laccaam-i k'armánco	'Lachaame's relative'
ʔantabaʔ-i ládo	'hen's feather'
meent-i horoóre	'hair of women'

Tadesse Sibamo describes the vowel **-i** as a linker in the last three examples above (transcribed as one word in the source).[170] Is it possibly a genitive marker or an epenthetic **-i-**, as it is also attested in the genitive constructions with the glottal stop (see the subsection on Genitive ʔ)?

In Gedeo, according to Eyob Kelemework (2015, pp. 335, 236), the order of the elements in the noun phrase is that "the head noun usually comes on the right-most side, and all the modifiers on its left. Nonetheless, this order of elements is flexible. Especially, if the modifiers are possessive pronouns, demonstratives, and numerals, the head noun may appear before its modifiers". For instance, the attributive nouns or adjectives may also follow the nouns they modify.

belt-i	manj-i-cco-ti=ke	me??-ø-e
child-nom.m	person-ep-sg-loc-gen.m	go-3sg.m-sg.perf

'The son of the man / woman went.'

[170] Tadesse Sibamo (2015) transcribes them as one word. However, they are most likely phrases (compare *maaʔlimíne* 'butchery' with Ritter's (2007) *laʔl mine* 'cattle area (in house)'. Similarly, the first three examples from Hudson).

manj-i-cco-ti=t'i	belto	me??-i-t-e
person-ep-sg-loc-gen.f.nom	child	go-ep-3sg.f-sg.pfv

'The daughter of the man / woman went.'

The above examples show that the genitive attributive singular noun *manjicco* 'person' follows the Modified noun *belto* 'child'. Such a flexible order is also found in Kambaata, but for noun phrases within a sentence (Treis, 2008, p. 102).

Treis (p.c; 2008, p. 102) and Kawachi (2007, p. 7, fn. 10) reject Hudson's (1976, pp. 254–255) claims where he posits that Kambaata and Sidaama ordinarily use apposition only and no morphemic marker to indicate the genitive. In Sidaama, Kawachi regards Hudson's example with the proper nouns *besʃa ooso* 'Besha's children', at least in the dialect of Sidaama he examined, as ungrammatical. It should instead be *besʃ-í ooso* 'Besha's children' with the genitive suffix **-i** of the masculine proper (and common) nouns ending in **-a**.

Gender

Except for in Hadiyya, gender distinction seems central to the genitive construction in HEC. In the Sidaama and Kambaata languages, gender refers to the gender of the possessor. This is in contrast to the genitive with the exponents **k** and **t** in Burji and Gedeo, where the choice of **k** (m) or **t** (f) depends on the gender of the possessed (see Hudson 1976, p. 254; Tesfaye Baye, 2015, p. 117; Eyob Kelemework, 2015, p. 115).

This split in gender assignment (possessor-controlled in Sidaama/Kambaata vs. possessed-controlled in Burji/Gedeo) is notable. If PHEC had a single consistent pattern, the current distribution would require at least one branch to have innovated a reversal. Alternatively, the genitive constructions in the two subgroups may have different historical origins, with Burji/Gedeo retaining a possessed-agreement pattern (comparable to Lowland East Cushitic) while PNHEC innovated a possessor-controlled vowel/suprafix system. This question requires further investigation.

In summary, several processes appear to have shaped the development of genitive markers of the HEC languages. Lenition is evident in the proposed shift from **k** to **h** and potentially to zero in Kambaata and Sidaama (§3.2.2). Assimilation is noted in Burji voicing assimilation (§3.1.1) and potentially in Kambaata vowel harmony within compounds. Coalescence/glottal stop elision

(§3.3.3) is suggested to explain Sidaama and Kambaata vowel/suprafix genitives, in which the stem and the vowels following the lost genitive consonant might have been influenced (§2.4). Epenthesis is proposed for the **-n-** element preceding consonants in Burji/Gedeo and reasonably for others, potentially arising after degemination (§3.4.3). These processes collectively illustrate the dynamic phonological changes influencing morphological forms in HEC.

4.1.4.4 Historical Development of the Genitive

There are two types of genitive construction. The first type of genitive is found without a following clitic; hence, it has no gender association with the head noun of the genitive construction in Hadiyya, Kambaata, and Sidaama. The second type includes the genitive suffix ***-ʔ**, followed by one of the gender clitics ***=k** (m), and ***=t** (f), which in turn are preceded by epenthetic **-n-** in the common nouns. Kambaata and Hadiyya show evidence of the first type only, while Sidaama has reflexes of both types, and Burji and Gedeo show the second type only.

The genitive ***-ʔ** is retained in Hadiyya in the form of fusion (§3.3.1) and metathesis (§3.3.2) depending on the sonority of the stem-final consonant. It is still evident in the glottalized consonant of the genitive clitic **=t't'** (§3.3.1) in Gedeo. The compensatory vowel doubling and accentuation **-V́(V)** (gen) (§2.2) in Kambaata and Sidaama is further evidence suggesting the loss of the following glottal stop of the genitive (see §2.4 and the subsection on Genitive ʔ).[171] Additionally, the gemination of the following consonant of the clitic **t** (f) and **k** (m) suggests a fusion of the **ʔ** (gen) and the following consonant. It is retained as such (**tt** (gen.f) and **kk** (gen.m)) on proper nouns and pronouns in Burji and Gedeo but was otherwise degeminated into a single consonant (§3.1.1). Consequently, the epenthetic **-n-** occupied the place of the first consonant, i.e., **n=t** (gen=f) and **n=k** (gen=m) upon degemination of **tt** (gen.f) and **kk** (gen.m) on Burji and Gedeo common nouns and Sidaama genitives with clitic **k** (m) and **t** (f) (§3.4.3). This means that all three languages share the fusion/assimilation stage, which yielded the geminates ***-kk** (gen.m) and ***-tt** (gen.f), and which in turn encode the older genitive and the clitic construction ***-ʔ=k** (gen=m) and ***-ʔ=t** (gen=f). Therefore, the current epenthetic **-n-** before the clitic on the common nouns holds the place of the

[171] Treis (2008, p. 37, fn. 23) reported a puzzling instance of a sequence **ʔ**r found on one of the (shared) lexemes where **ʔ** cannot be analyzed as a middle marker: in Kambaata *gorʔá* 'green, unripe berry' and Sidaama *goʔra* 'wild berry similar to blackberry and raspberry' (cited from Kawachi 2007, p. 29, fn. 1).

old genitive ***-ʔ** in the genitive construction (i.e., ***n=k** (gen=m) and ***n=t** (gen=f)). Below is a summary of these processes as attested across the different languages (Assim. = assimilation (1 first, 2 second assimilation), Gem. = gemination, Degem. = degemination, Epen. = epenthesis, Meta. = metathesis).

	Genitive + final/Clitic	Assim.1 + Gem.	Fusion + Gem.	Degem. / Epen.	Assim.2 /Meta.
Bur, Ged, Sid	***ʔ + *k →**	**kk**	**-**	**nk**	**nn**
Bur, Sid	***ʔ + *t →**	**tt**	**-**	**nt**	**nn**
Ged, Had	***t + *ʔ →**	**-**	**t't' (← *t+ʔ)**	**nC'**	**-**
Had	***$C_{i[-son]}$ + *ʔ →**	**-**	**C'C'**	**-**	**-**
Had	***$C_{i[+son]}$ + *ʔ →**	**-**	**-**	**-**	**ʔC**
Sid, Kam	***V_i + *ʔ →**	**-**	**-**	**V́(V)**	**-**

This schema proposes the diachronic pathways but should be read with caution: Not all intermediate stages are directly attested, and the ordering of processes is inferred from synchronic distributions. The core claim is that PHEC ***-ʔ** combined with clitics ***=k** (m) and ***=t** (f) and subsequently underwent assimilatory/fusion processes. It is an arguable interpretation based on the independent evidence from Hadiyya glottalization, Gedeo glottalized **t't'**, and Kambaata/Sidaama compensatory doubling/accentuation. The **-n-** epenthesis is best understood as a repair strategy following degemination, filling the position vacated by the lost first geminate component.

A second assimilation (Assim.2) ***nk** / ***nt** → **nn** is unique to Sidaama.[172] Additionally, Sidaama has a following **-i-**, which is inserted in **-nn-i-** to avoid the **CCC** cluster that would result when a following consonant-initial morpheme (e.g., clitic) is attached.

Almost all the retained forms of the genitive construction show evidence of the glottal stop. Treis (2008, p. 226; fn. 134) notes that in Kambaata, the linker **n** might also be called a grammatically determined epenthetic consonant. Although that is not applicable to the Kambaata genitive, historically, it aligns with the morphophonological epenthesis process of **-n-** (§3.4.3) after the (***-kk-**, ***-tt-**, ***-t't'-**) degemination and may apply to most, if not all, the instances of **-n-** presented above on the common nouns in the genitive

[172] The assimilation patterns 1 and 2 are not improbable (§3.1.1); for instance, Gedeo **ʔt** assimilating to **tt** in *mado?-* 'wound' → *madot-t-e* (wound-3sg.f.perf) (Eyob Kelemework, 2015, p. 62). However, it is unclear whether all languages have synchronic morphophonemic evidence or whether we must rely solely on diachronic explanations. The same applies for Sidaama ***nk** / ***nt** → **nn**.

construction. The elements involved in that process are the genitive ***-ʔ**, stem-final $\mathbf{C_i}$, $\mathbf{V_i}$, and the clitic ***k**/***t** (Bur = Burji).

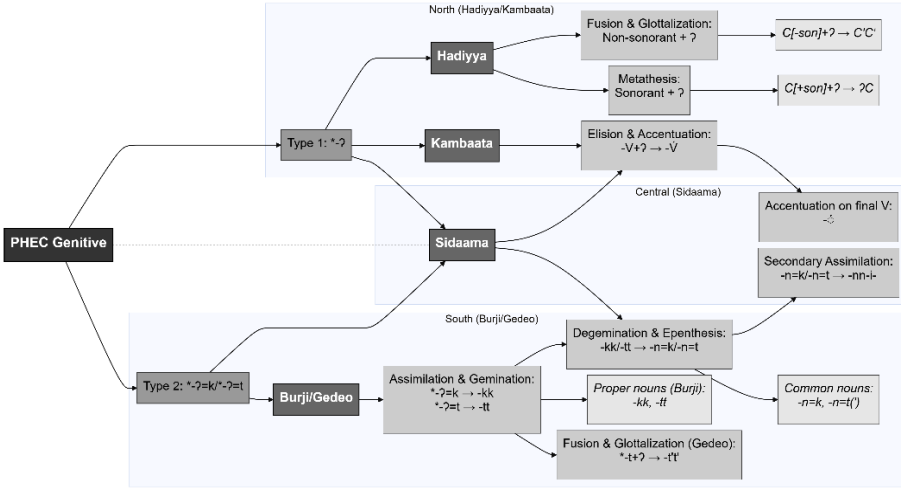

Figure 3: *Historical Development of the Genitive*

The comparative data reveal a typological and geographical cline in HEC genitive constructions: from the morphologically complex systems of Burji and Gedeo in the south through the central mixed system of Sidaama, to the simple system of Hadiyya and Kambaata in the north. Burji and Gedeo present the most complex systems in a robust paradigm that distinguishes case (nominative/accusative/predicative), gender (masculine/feminine), and noun type (common/proper). The variation in these exponents' retention and function reflects broader language change processes in the HEC languages, including morphological restructuring, functional reanalysis, and phonological attrition.[173]

The above reconstructions are considered arguable proposals because the diversity in the genitives might alternatively suggest independent innovations in each language. However, such an alternative interpretation ignores the comparable parallels in Burji and Gedeo, on the one hand, and Kambaata and

[173] This reconstruction, particularly the proposed origin of the **k**/**t** clitics from copulas in cleft constructions leading to a word order shift, aligns with the analysis by Tosco (1994). It differs from Hetzron's (1980) hypothesis, which, while acknowledging the **k**/**t** agreement as an important link between HEC and Agaw, suggested these markers were transferred from the possessed noun to the possessor. Tosco argues that the cleft-origin model better accounts for the attested structures without invoking such a transfer.

Sidaama, on the other hand. Therefore, the proposed reconstructions could be solidified upon further research into individual languages (especially Hadiyya), more harmonized analysis, and use of the terminology where languages exhibit comparable systems.

4.1.5 Gender and Case

HEC nominal morphology treats gender and case as formally intertwined categories. The masculine marker ***-k** and feminine marker ***-t** recur across demonstratives, definiteness markers, and pronominal endings; the case vowels ***-i** (nominative), ***-o** (masculine accusative), and ***-e** (feminine accusative) attach directly to these gender exponents or appear fossilized in pronominal paradigms. This section surveys the distribution of both markers and evaluates what can be reconstructed to PHEC or even Pre-PHEC, beginning with gender (§4.1.5.1) and then turning to case (§4.1.5.2).

4.1.5.1 Gender

The masculine marker ***-k** and feminine marker ***-t** are pan-Cushitic exponents that recur across the HEC languages in demonstratives, definiteness markers, case suffixes, and agreement morphology. Their distribution supports reconstruction to PHEC with secure confidence; the individual languages differ only in which grammatical functions these consonants continue to serve productively.[174]

[174] Data sources: Burji (Tesfaye Baye, 2015, pp. 100, 104), Hadiyya (Hudson, 1976, p. 251; Tadesse Sibamo, 2015, p. 53), Kambaata (Treis, 2008, pp. 108, 129–130, 361), Sidaama (Kawachi, 2007, pp. 200, 483), Gedeo (Dawit Tilahun, 2019, pp. 91, 112; Gasparini, 1995, pp. 9, 2).

	-k (m)	-t (f)	Function
Burji	**-ʃi̱**	**-tteé**	gender
	kú	**-cí**	demonstrative ‘this’
	-gu	**-ʃi**	subject definite
Had	***-k**	**-ta**	gender
	kúki	**túte**	definiteness demonstrative ‘this’
Kam	**-ha, -hu**	**-ta, -ti**	secondary case / gender
	ka, ku	**ta, ti**	demonstrative attribute ‘this’
Sid	**ko**	**te**	adnominal demonstrative ‘this’
	-ho	**-te**	dative case (Unmodified common nouns)
Ged	**koo**	**tee**	vocative (proper nouns)
	kunni	**tinni**	demonstrative pronoun ‘this’
	-ki	**-ti**	definite (article)
	-k-eʔe	**-t-eʔe**	dative case (common nouns)

Hadiyya **-ta** is found on some nouns (Hudson, 1976, p. 251), while ***-k** is only recoverable from underlying forms via a morphophonological process triggered by the attachment of some suffixes (§3.3.4). In Kambaata, the vowels following the gender exponents are case markers: **-a** for the accusative and oblique and **-i** (f) or **-u** (m) for the nominative (Treis, 2008, pp. 127, 322). Treis (2008, p. 108) analyzes the development of the above functions in Kambaata as a “recent grammaticalization of case- and gender-sensitive demonstratives into secondary case / gender markers”. Yri & Pepper (2019) specify that, in Sidaama, the vowels **-u** and **-o** follow the masculine gender consonant **k-** and the vowels **-i** and **-e** follow the corresponding consonant **t** of the feminine in the demonstrative pronouns, but they do not associate them with a specific case as in the other HEC languages. The Gedeo definite markers **-ki** and **-ti** become **-ka**, **-ta** in the accusative and **-ke**, **-te** in the other cases (Gasparini, 1995, p. 2).

Hudson (1976, p. 251) lists some of the occasions when the gender exponents occur on the HEC nouns: (a) in their agreement with verbs, (b) in their choice of demonstratives, (c) in their construction of the genitive (in some languages, §4.1.4.1), and (d) in their choice of a suffix (article), to some extent.

To summarize: PHEC ***-k** (masculine) and ***-t** (feminine) are securely reconstructed on the basis of uniform attestation across demonstratives, definiteness markers, and dative/genitive case morphology. The consonants themselves are stable; the vowels that follow them encode case and other grammatical distinctions, which are addressed in §4.1.5.2. The pan-Cushitic

attestation of ***-k** / ***-t** (Tosco, 2020, p. 294; Hudson, 1976, p. 246) situates the HEC pattern within the broader Afroasiatic typology of gender marking.

4.1.5.2 Case

The terms nominative and accusative (some sources refer to it as the absolutive) refer primarily to form; subject and object denote function. This distinction matters because the pronominal paradigms preserve case marking that is no longer productive in standard nominal inflection (§1.5.2).

As an alternative to the null hypothesis that the current terminal vowels have nothing to do with any fossilized case marking, one could argue for potential old case functions that the TVs used to serve, based on remnants of synchronic evidence (see §2.2.3).

Accusative

The accusative case markers, reconstructed as ***-o** (← ***-ó**) for masculine and ***-e** (← ***-é**) for feminine, survive primarily in fossilized pronominal forms rather than in productive nominal paradigms. Their distribution is most clearly visible in the third-person singular object pronouns and in the interrogative. The following subsection presents the evidence for these reconstructions and notes where individual languages have generalized or restructured the inherited pattern.

Accusative *-e and *-o

The best evidence for the accusative (**-e** (f) and **-o** (m)) vowels comes from the singular personal pronouns. The fossilized feminine accusative ***-e** is reconstructed in the pronouns [151] **is-e* ‘her’ and [28] **ay-e* ‘who?’. The nominative counterparts are [6] **is-i* ‘she’ and [28] **ay-i* ‘who?’ (→ Kam and Sid *ayi* (nom) ‘who?’).

	[151]		**[28]**		**[292]**	
PHEC	***ise**	‘her’	***aye**	‘who?’	***iso**	‘him’
Burji	**ʔiʃeé**	‘her’	**ayyee**	‘who?’	**ʔisi̯**	‘him’
Had	**ʔíse̯**	‘her’	**ayye**	‘who?’	**it’t’o**	‘him’
Kam	**isé-ta**	‘her’	**aye**	‘who?’	**isú**	‘him’
Sid	**ʔisé**	‘her’	**aye**	‘who?’	**ʔisó**	‘him’
Ged	**isee, iseeʔe**	‘her’	**ayye**	‘who?’	**isoo, isooʔo**	‘him’

The masculine counterpart **-o** of the feminine accusative **-e** above is also inferred as an arguably fossilized suffix from the pronoun [292] PHEC **is-o* 'him'. Kambaata *isú* 'him' shows a regular **-u**. In Burji, the vowel **-o** of the object pronoun has been restructured due to analogy with the vowel of the subject pronouns [7] *ʔísi̯* 'he, it' and [6] *ʔíʃi̯* 'she', thus becoming *ʔisi̯* 'him'.

In set [407], **ebel-, *ebal-* 'so-and-so' behaves similarly to pronouns with the TVs **-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* and Kambaata *ebale* 'so-and-so', even though Kambaata **-o** is an irregular reflex as noted in the terminal vowel **-o** section in §2.2.3.1.

In addition to the pronouns, the demonstratives (pronouns/determiners/adjectives) provide fossilized evidence of the arguably proposed accusative ***-o** and ***-e**.

		Masculine Acc	**Feminine Acc**	**Gloss**
Burji	[418]	**ka**	**ta**	'this'
Had		**ka**	**ta**	'this' (det)
Kam		**ka**	**ta**	'this' (det)
Sid		**ko, konne, koe**	**te, tenne**	'this'
Ged		**konne**	**tenne**	'this' (det)
Burji	[281]	**ci**	**ci**	'these'
Had		**kakeen**	**kakeen**	'these'
Kam		**kara**	**-**	'these' (pron)
Sid		**konne, ko(ʔo)**	**tenne, te(ʔe)**	'these'
Ged		**konneera**	**tenneera**	'these'

The TV is mainly **-a** in Burji, Hadiyya, and Kambaata. However, Sidaama and Gedeo may have maintained a distinction in vowel quality with **-o** for the masculine and **-e** for the feminine before the attachment of **-nnV(V)**.

The singular demonstrative [418] can be reconstructed as the masculine accusative ***k-o** and feminine accusative ***t-e** (f.acc) 'this'. This reconstruction assumes that the Gedeo and Sidaama forms have preserved a more archaic vowel contrast **-o** (m) and **-e** (f). However, such an assumption disregards the fact that the gender distinction is already realized by the consonants ***k** (m) and ***t** (f), making the gender distinction that the vowels are supposed to realize in addition to the case marking redundant. Therefore, an alternative reconstruction of the accusative vowel would be accusative case ***-a** based

on the Burji, Hadiyya, and Kambaata reflexes, with the gender realized via the consonant contrast, i.e., ***ka** (m.acc) and ***ta** (f.acc).[175] One might argue in favor of the first reconstruction that the consonantal gender distinction of the demonstratives might have encouraged the later neutralization of the vowel distinction, or that the vowel **-a** is a result of vowel coalescence: **ko** and **te** + **a** → **ka** and **ta** in Burji, Hadiyya, and Kambaata. The pronouns provide further evidence, in support of the first reconstruction, for maintaining ***-e** (f) and ***-o** (m) in the accusative.

The following examples are identical to the ones presented under the section on variation in terminal vowels (2.2.3.2). However, these examples include the respective gender of the stem in each language to examine if the TV of the nouns matches the proposed genders (gen) in the accusative, assuming that some might be fossilized case markers:

	PHEC	Gloss	Had	Sid	Ged	Kam	Burji
[174]	*hamas-	'snake'	**-aʃʃa**	**-o**	**-me**	**-u**	**-i**
			-	(m)	(f)	(m)	(m)
[137]	*lal-	'cattle'	**-o**	**-o**	**-o, -i**	**-u**	**-i**
			-	(m)	(m)	(m)	(m)
[444]	*bagaz-	'spear'	**-o**	**-o**	**-o**	**-u**	**-i**
			-	(m)	(m)	(m)	(m)
[95]	*ʃomb-	'lung'	**-**	**-o**	**-o**	**-**	**-i**
			-	(m)	(m)	-	(m)
[606]	*gubeed-	'thigh'	**-o**	**-a**	**-a**	**-**	**-i**
			-	(m)	-	-	-
[9]	*arrab-	'tongue'	**-o**	**-o**	**-o**	**-i-ta**	**-a**
			-	(m)	(m)	(f)	(f/m)
[4]	*gulib-, *gilub-	'knee'	***-co**	**-co, -e**	**-o**	**-i-ta**	**-a**
			-	(m)	(m)	(f)	(f)
[382]	*oll-	'neighbour'	**-aʔa**	**-a**	**-aʔa**	**-oo**	**-a**
			(pl)	(m)	(m)	(m)	(m/f)
[51]	*lokk-, *lekk-	'foot, leg'	**-o**	**-a**	**-a**	**-a-ta**	**-a**
			-	(f)	(f)	(f)	(f)
[293]	*waʔ-	'water'	**-o**	**-a**	**-aʔe, -eʔe**	**-a**	**-a**
			-	(m)	(m)	(m)	(m)
[88]	*gudum-	'shoulder'	**-o**	**-**	**-o**	**-u**	**-a**
			-	-	(f)	(m)	(f/m)
[195]	*fal-, *falfal-	'omen, magic'	**-o**	**-**	**-o**	**-a**	**-aa**
			-	-	(f)	(m)	(m)
[5]	*san-	'nose'	**-e**	**-o**	**-o**	**-u-ta**	**-a**

[175] Blažek & Malášková (2016) identify the final vowel **-a** on the first-person object pronouns as the absolute case suffix: Oromo Tulema *ánā*, *na*, Konso *ána*, and Mashile *ana*. As mentioned above, "absolutive" is used interchangeably with "accusative" in the HEC literature.

	PHEC	Gloss	Had	Sid	Ged	Kam	Burji
			-	(f)	(f)	(f)	(f)
[439]	*reħ-	‘death’	**-o**	**-o**	**-o**	**-u-t**	**-a**
			-	(f)	(f)	(f)	(f)
[206]	*awd-	‘threshing floor’	**-a**	**-umme**	**-e**	**-u**	**-a**
			-	(f)	(f)	(m)	(f)
[408]	*mut-	‘awl for basketwork’	**-a**	**-C’C’e**	**-a**	**-a**	**-a**
			-	(f)	-	(m)	-
[381]	*k’ul-, *k’ul~k’ul~l-ʔ-	‘leech’	**-a**	**-e**	**-a**	**-a**	**-ee**
			(pl)	(f)	-	-	(m)
[1]	*ill-	‘eye’	**-e**	**-e**	**-e**	**-i-ta**	**-a**
			-	(f)	(f)	(f)	(f)
[415]	*afal-	‘liver’	**-e**	**-e**	**-**	**-i-ta**	**-a**
			-	(f)	-	(f)	(f)
[54]	*min-	‘house’	**-e**	**-e**	**-e**	**-i**	**-a**
			-	(m)	(m)	(m)	(m)
[550]	*ibb-	‘louse’	**-a**	**-e, -iicco**	**-e**	**-i-ta**	**-a**
			-	(f)	(f)	(f)	(f)
[224]	*t’ur-	‘dirt of body, clothing’	**-a**	**-e**	**-e**	**-i**	**-ee**
			-	(m)	(m)	(m)	(f)
[544]	*c’ub- (PNHEC)	‘dagger’	**-eʔe**	**-e, -a**	**-e**	**-i-ta**	
			-	(f)	-	(f)	

The above examples show the respective gender (masculine (m), feminine (f)) underneath the terminal vowel of each noun. Those with only a dash “**-**” indicate absence of gender identification in the source (e.g., all Hadiyya examples) or lack of a cognate. The available Hadiyya sources do not indicate the gender of the nouns. Kambaata **-ta** is the feminine accusative suffix, and it is likely added after the TV (mostly **-i** and **-u**). Some examples include additional suffixes, such as the singulative or derivational suffix in [174], [550], and [4]. 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 in the same language, e.g., Sidaama **-e, -a** in [544], and Gedeo **-o, -i** in [137], each is from a different data source.

The following paragraphs address the question of whether there is any association between the quality of the terminal vowel and the assigned gender (i.e., a fossilized marker) or whether the gender is inherent to the root.

Concerning the first half of the question, the data show that a single stem can end in **-a**, **-i**, **-o** (Kambaata **-u**), or **-e**, depending on the language, e.g., [606]

Hadiyya **-o**, Gedeo and Sidaama **-a**, Burji **-i**. In general, the observed NHEC patterns in terminal vowel quality and gender match, including masculine (m) forms that show a back TV (**-o**), while feminine forms surface with a front TV (**-e**). The corresponding vowels in Kambaata are **-u** and **-i**, respectively. In contrast to Kambaata, Burji shows **-i** for the masculine. The TV **-a** appears on both masculine and feminine stems in all HEC languages. The following examples show regular correspondences aligning with these patterns for the masculine and feminine found in 8 out of the total 23 examples:

	Had		**Ged**		**Sid**		**Kam**		**Burji**	
[95]	**-**	-	**-o**	(m)	**-o**	(m)	**-**	-	**-i**	(m)
[174]	**-**	-	**-**	-	**-o**	(m)	(m)	(m)	**-i**	(m)
[444]	**-o**	-	**-o**	(m)	**-o**	(m)	**-u**	(m)	**-i**	(m)
[137]	**-o**	-	**-o, (-i)**	(m)	**-o**	(m)	**-u**	(m)	**-i**	(m)
[415]	**-e**	-	**-**	-	**-e**	(f)	**-i-ta**	(f)	**-a**	(f)
[1]	**-e**	-	**-e**	(f)	**-e**	(f)	**-i-ta**	(f)	**-a**	(f)
[550]	**(-a)**	-	**-e**	f	**-e**	(f)	**-i-ta**	(f)	**-a**	(f)
[544]	**-eʔe**	-	**-e**	f	**-e, (-a)**	(f)	**-i-ta**	(f)	**-**	-

Overall, Kambaata exhibits the highest degree of gender and vowel quality matching with **-i** in the feminine and **-u** in the masculine stems in the above and other examples.

However, there is counterevidence for the above patterns, in which back vowels occur in feminine stems and front vowels occur in masculine stems.

	Had		**Ged**		**Sid**		**Kam**		**Burji**	
[224]	**-a**	-	**-e**	(m)	**-e**	m	**-i**	(m)	**-ee**	(f)
[54]	**-e**	-	**-e**	(m)	**-e**	m	**-i**	(m)	**-a**	(m)
[439]	**-o**	-	**-o**	(f)	**-o**	(f)	**-u-t**	(f)	**-a**	(f)
[5]	**-e**	-	**-o**	(f)	**-o**	(f)	**-u-ta**	(f)	**-a**	(f)

As mentioned above, the low TV **-a** is found in both masculine and feminine stems (see §4.1.5.1).

	Had		Ged		Sid		Kam		Burji	
[51]	**(-o)**	-	**-a**	f	**-a**	f	**-a-ta**	f	**-a**	f
[293]	**(-o)**	-	**-aʔe, -eʔe**	m	**-a**	m	**-a**	m	**-a**	m
[382]	**-aʔa**	pl	**-aʔa**	m	**-a**	m	**(-oo)**	m	**-a**	m/f
[408]	**-a**	-	**-a**	-	**-**	-	**-a**	m	**-a**	-
[381]	**-a**	pl	**-a**	-	**(-e)**	f	**-a**	-	**(-ee)**	m

The TVs that do not fit the quality of the ascribed pattern are placed between parentheses, e.g., **-e** in [381], **-o** in [51], and **-a** in the previous examples [550] and [544].

Overall, the terminal vowels of the PHEC accusative pronouns **is-e* (3sg.f.acc) ‘her’ and **is-o* (3sg.m.acc) ‘him’ coincide with the high frequency of terminal **-o** and **-e** in the absolutive/accusative nouns in at least Gedeo, Hadiyya, and Sidaama. They are also the only feature of the PNHEC noun root **ar-* that distinguishes gender and sex between [212] **ar-o* (m) ‘husband’ and [271] **ar-e* (f) ‘wife’.

*Accusative *-é and *-ó*

Accented ***-é** (f.acc) and ***-ó** (m.acc) could be reconstructed for the PHEC accusative case as the synchronic evidence proposes. Prior to the current accented single vowels, the accusative might have been double ***-ée** (f.acc) and ***-óo** (m.acc) or vowel and consonant sequences of the shape ***-V́C**, representing a valid accent-bearing syllable, as discussed in §2.2.

The alternative scenario (null hypothesis) is that the gender is inherited from the root. Such a scenario requires regularity in the assigned gender of each cognate set, regardless of the quality of the TV in the individual languages. Half of the 23 examples listed above have the same gender across all languages in which they are attested, regardless of TV.

	PHEC	Gloss	Gender
[206]	***awd-**	‘threshing floor’	(f)
[550]	***ibb-**	‘louse’	(f)
[415]	***afal-**	‘liver’	(f)
[1]	***ill-**	‘eye’	(f)
[4]	***gulib-**	‘knee’	(f)
[51]	***lokk-, *lekk-**	‘foot, leg’	(f)
[439]	***reħ-**	‘death’	(f)
[54]	***min-**	‘house’	(m)
[293]	***waʔ-**	‘water’	(m)
[95]	***ʃomb-**	‘lung’	(m)
[137]	***lal-**	‘cattle’	(m)
[444]	***bagaz-**	‘spear’	(m)
[382]	***oll-**	‘neighbour’	(m)
[224]	***t’ur-**	‘dirt of body, clothing’	(m)

There are two exceptions in the Burji examples, where [382] is assigned masculine in one source and feminine in another, and where [224] is feminine in contrast to its cognates in the rest of HEC.

In summary, the immediate accusative reconstruction posits ***-e** (f.acc) and ***-o** (m.acc) based on the surface forms of third-person pronouns and the terminal vowel patterns in nouns. The more abstract layer (accented ***-é** and ***-ó**) accounts for phontactic properties (accent position, syllable weight) that the unaccented single vowels alone do not explain. Therefore, they are traceable to earlier double vowels or vowel-consonant sequences pending broader East Cushitic confirmation. These competing hypotheses do not affect the core reconstruction, which is arguable at the PHEC level.

Nominative

The nominative case markers, reconstructed as ***-i** for feminine and ***-u** for masculine, are preserved in third-person subject pronouns and demonstratives. Unlike the accusative, where the vowel contrast is clearly attested across most HEC languages, the nominative markers show considerable leveling: many languages generalize ***-i** to both genders, obscuring the original distinction. The evidence for ***-u** as the masculine nominative is strongest in Burji, Kambaata, and Sidaama demonstratives;

elsewhere, it has been replaced by analogy. The following subsections present the immediate evidence for ***-i** and ***-u**, then propose a more abstract reconstruction involving glides ***-y** and ***-w** to account for phonological properties such as devoicing and lack of accent.

Nominative *-i and *-u

Terminal **-i** is most frequently found on nominative pronominals in all HEC languages: Gedeo (Gasparini, 1995, p. 1), Hadiyya (Tadesse Sibamo, 2015, p. 59), Burji (Tesfaye Baye, 2015, pp. 48, 104, 115), Sidaama (Anbessa Teferra, 2014, p. 31; Indrias Xa'miso et al., 2007, p. 637), Kambaata (Treis, 2008, pp. 322, 130, 127). Thus, nominative ***-i** is arguably reconstructable to PHEC. It is most likely the feminine nominative marker in the third-person subject pronouns, e.g., PHEC [6] **isi* 'she', equivalent to the feminine accusative ***-e** in PHEC [151] **ise* 'her'.

The back counterpart **-u** is found as the terminal vowel of nominative forms in various nominals in Burji (Tesfaye Baye, 2015, pp. 48, 104, 115), Kambaata (Treis, 2008, pp. 322, 130, 127), and Sidaama (Anbessa Teferra, 2014, p. 31; Indrias Xa'miso et al., 2007, p. 637) but never attested in Gedeo or Hadiyya. It alternates with **-i** in Sidaama. In Burji, Kambaata, and Sidaama, there is a tendency to associate **-u** with the masculine nominative in terms of gender. Thus, it is arguably reconstructed as a remnant marker of (pre-)PHEC ***-u** (m.nom) equivalent to the feminine ***-i** (f.nom). However, the latter has likely been generalized by analogy, resulting in frequent attestation of **-i** in both the masculine and feminine nominative.

The demonstratives (pronouns/determiners/adjectives) provide the best fossilized evidence of the tentatively proposed nominative masculine ***-u** and arguable feminine ***-i**, among the pronominals.[176]

[176] Sources of additional examples (singular and plural): Sidaama (Kawachi, 2007, p. 202), Burji (Hudson, 1976, p. 256), Gedeo (Dawit Tilahun, 2019, p. 143).

		Masculine Nom	Feminine Nom	
[418]	Burji	**ku**	**ci**	‘this’ (det)
	Had	**ku(k)**	**tu(t)**	‘this’ (det)
	Kam	**ku**	**ti**	‘this’ (det)
	Sid	**ko, kuni, kuyi**	**te, tini, tiyi**	‘this’
	Ged	**kunni**	**tinni**	‘this’ (det)
[281]	Burji	**kan(d)’u-gu**	**cin(d)’i-ʃi**	‘these’ (pron)
	Had	**kukeen**	**tuteen, tu keen**	‘these’
	Kam	**kuru**	**-**	‘these’ (pron)
	Sid	**kuni, kuri**	**tini, kuri**	‘these’
	Ged	**kunniiri**	**tinniiri**	‘these’

The singular demonstrative [418] can be reconstructed as masculine nominative ***k-u** (m.nom) and feminine nominative ***t-i** (f.nom) ‘this’. Such a reconstruction is supported by the reflexes of Burji, Kambaata, and one of the Sidaama variants. The vowel distinction leveling to **-u** in Hadiyya might be due to the redundancy of gender assignment via the consonants ***k** (m) and ***t** (f) and the vowels ***-u** (m) and ***-i** (f) at the same time. The same applies to the Sidaama and Gedeo variants, with **-i** being generalized after the addition of the nasal **-n** (or glide **-y**).

The plural counterparts of the demonstratives in [281] are provided to give more context. They also help explain the nasal **-n** in the Sidaama and Gedeo singular forms, which are a generalization of the plurative forms (see plurative ***-n** in §4.1.1.2).

To summarize the immediate reconstruction: PHEC ***-i** (f.nom) is attested across all HEC languages, appearing on third-person feminine subject pronouns, e.g., [6] **is-i* ‘she’) and generalized to other nominative contexts. PHEC ***-u** (m.nom) is more restricted and preserved in masculine demonstratives, e.g., [418] **k-u* ‘this’ (m) and some masculine nominative pronouns in Burji, Kambaata, and Sidaama, but leveled to ***-i** in most other cells.

*Nominative *-y and *-w*

The nominative vowels are widely reported as unaccented and devoiced (§2.2.3), even sometimes inaudible in comparison to the accusative. This could be addressed with further abstraction, suggesting that instead of or prior

to the vowels ***-i** (f) and ***-u** (m), there might have been consonant glides ***-y** (f) and ***-w** (m) marking the third-person singular nominative.

	[6]		[7]		[418]	
PHEC	***is-y** → ***is-i**	‘she’	***is-w** → ***is-u**	‘he’	***k-w** → ***k-u**	‘this’ (m)
Burji	**ʔíǰ̣**	‘she’	**ʔísị̣**	‘he, it’	**ku**	‘this’ (det)
Had	**ʔísị̣**	‘she’	**isse, issoʔi, it’t’o**	‘he/she’ (honorific), ‘he’	**ku(k)**	‘this’ (det)
Kam	**íse**	‘she’	**ís, ísu**	‘he’	**ku**	‘this’ (det)
Sid	**ise**	‘she’	**ʔísi**	‘he’	**ko, kuni**	‘this’
Ged	**ise**	‘she’	**isi**	‘he’	**kunni**	‘this’ (det)

In the nominative, the HEC languages generalized the ***-y** from the feminine to the masculine. Kambaata *isu* ‘he’ has a terminal **-u** due to analogy with the accusative counterpart *ísu* ‘him’. However, almost all the reflexes of the masculine nominative demonstrative have retained the expected vowel quality **-u**.

The third-person plural adds ***-n** to the singular stem.

	[11]	
PHEC	***is-n-**	‘they’ (subj)
Burji	**ʔisínụ̣**	‘they’ (subj)
Had	**it’t’uw, issuw**	‘they’ (subj)
Kam	**issoʔóot**	‘they’ (subj)
Sid	**ʔínsa**	‘they’ (subj)
Ged	**insaʔnne, insaʔne**	‘they’ (subj)

The reconstruction of the plural above is tentative due to variation and innovations across the individual languages. There is no gender distinction synchronically in the third-person plural.[177]

The further abstraction proposes pre-PHEC ***-y** (f.nom) and ***-w** (m.nom) as consonant glides that later vocalized to ***-i** and ***-u**, respectively. This layer accounts for several otherwise puzzling features: (a) the characteristic devoicing of nominative vowels across HEC (voiced consonants are less sonorous than vowels, yielding devoiced reflexes upon vocalization); (b) the absence of accent on nominative terminal vowels (the earlier glides were

[177] The glide **-w** found in Haddiyya *it’t’uw*, *issuw* ‘they’ (subj) might be a plurative suffix (§4.1.1.2) rather than the proposed accusative **-w**.

consonants and thus not able to be accented); (c) the palatalized reflexes **ʃ** and **c** in Burji, explainable as the fusion of ***k** + ***y**, ***t** + ***y** (§§3.2.1, 3.1.2.1), and ***s** + ***y**;[178] and (d) parallels with other East Cushitic languages that preserve ***-y** in cognate pronominal forms. The glide hypothesis does not replace the immediate reconstruction but extends it to an earlier stage, providing a principled explanation for the phonological asymmetry between accusative (accented, voiced) and nominative (unaccented, devoiced) markers.

The TV of the nominative, ***-i**, has not only been generalized to both masculine and feminine but also to the other singular subject pronouns [2] **an-i* 'I' and [27] **at-i* 'you'. Therefore, Hudson (2006, p. 296) notes that the HEC languages generally mark nominatives with **-i**, while acknowledging that case marking is complex, is perhaps undergoing change, and needs more research. Beyond HEC, according to Blažek and Malášková (2016), the terminal vowels of the first-person PEC form **ʔani* and the variant **ʔanu*, found in Saho-Afar and Dullay (cf. also Beja *anū*), are noun subject case markers.

In concordance with the above proposed case markers, Hudson (1976, p. 253) reports that in Burji, the "alternation absolute form final e / nominative i, absolute form final o / nominative u is completely regular in proper names […]. Nouns in the accusative case appear in the absolute form, usually final o, a, or e."[179] This suggests a preservation of fully contrasting vowel reflexes of all the reconstructed case exponents in the proper nouns in Burji.

According to Treis (2008, p. 108), "Sasse's observation (1984:111) that the distinction between subject and object is often abandoned in Cushitic languages due to the interference with other grammatical categories (gender and focus) does not hold for Kambaata." The other grammatical categories Sasse mentioned have complicated the synchronic reality, but Treis's statement is accurate at least with regard to HEC, as shown in the above discussion.

[178] For example, in Burji [27] **at-y* becomes *ʔáʃi̯* 'you' (2sg.nom), [418] **t-y* becomes *ci* 'this' (sg.f.det), and [6] **is-y* 'she' becomes **iʃi̯*, i.e., the **-y** triggers the palatalization instead of the front vowel **-i**.
[179] In Konso, only names and pronouns show case (Ongaye Oda, 2013, p. 133).

If one discards all the above proposals of old functions, the terminal vowels on the pronominals have to be interpreted as having nothing to do with case marking (null hypothesis, see §2.2.3).

The case markers can be reconstructed at two levels. At the immediate PHEC level, based on direct surface correspondences, the accusative is marked by ***-e** (feminine) and **-o** (masculine), and the nominative by ***-i** (feminine) and ***-u** (masculine). These reconstructions rest on evidence from third-person pronouns, demonstratives, and terminal vowel patterns in nouns. However, several anomalies complicate this picture: Burji shows paradigm leveling toward ***-i** in both nominative and accusative masculine forms; Kambaata reflects ***-o** as ***-u** (a regular sound change) and ***-e** as ***-i** in some contexts; and the nominative ***-u** is widely replaced by generalized ***-i** outside of demonstratives.

A more abstract analysis can relate the case-vowel asymmetry to earlier phonological material. One possibility is that the accusative vowels continue an older accented opposition (**-é** for feminine vs. **-ó** for masculine), or derive from earlier VV/VC sequences. The nominative vowels are also compatible with an origin in pre-PHEC glides (**-y** for feminine and **-w** for masculine), which would account for their devoicing, lack of accent, and the palatalized consonants found in Burji demonstratives and pronouns. These abstractions do not replace the immediate PHEC reconstructions; they provide a possible account of why accusative exponents are accented and fully voiced while nominatives are typically unaccented and devoiced.

4.1.6 Personal Pronouns

Personal pronouns in Highland East Cushitic distinguish between nominative and accusative cases in all five languages. The comparison isolates recurring pronominal stems and segmentable case exponents, reconstructing paradigmatic structure only where correspondences are stable. Accusative forms (§4.1.6.1) show closer cross-language agreement than nominative forms (§4.1.6.2), especially in the plural, where Burji diverges from the NHEC languages in both stem shape and case-vowel reflexes. Across both paradigms, plural pronouns are typically marked by the plurative ***-n** (§4.1.1.2).

The paradigm distinguishes first, second, and third-person in singular and plural; only the third-person singular shows an obligatory gender contrast. The singular series is largely stable and supports reconstruction to PHEC with

limited ambiguity. The plural series shows extensive generalization (analogical leveling of stems and/or terminal vowels across cells) and paradigm restructuring (reallocation of inherited material across case/person slots).

4.1.6.1 Accusative Pronouns

The accusative (object) pronoun paradigm provides the clearest evidence for reconstructing PHEC case vowels. Singular forms show high cross-language regularity and support reconstruction of the main stems (**ey* 'me', **ke* 'you' (sg), **is-* '3rd person') together with their accusative exponents. Plural accusatives consistently retain the nasal plurative ***-n**, but their additional material varies by language and is treated here as later paradigm restructuring (i.e., language-specific reallocation or extension beyond the shared **-n** base). The generalized forms (i.e., leveled stems) are placed between parentheses. The following lists present the attested reflexes alongside PHEC and PNHEC reconstructions:

First-Person Object

	[103]		[197]	
PHEC	***e-y**	'me'	***n-y**	'us'
Burji	**ʔeé**	'me'	**niísi̯**	'us'
PNHEC	***e-y**	'me'	***n-y**	'us'
Had	**ʔeésẹ̯**	'me'	**neésẹ̯**	'us'
Kam	**ées**	'me'	**nées**	'us'
Sid	**ʔané**	'me'	**ninké**	'us '
Ged	**anee/aneʔe**	'me'	**(noʔooʔo)**	'us'

The first-person accusative pronouns are reconstructed with the **-y** shared by the other first-person pronouns (§4.1.4.2). Much of the diversity among surface cognates reflects morphological combination rather than stem replacement. Burji preserves a stem *ee* (on glide loss rule iii(b), see §2.4 on accent placement). In Hadiyya and Kambaata, *ee* coalesces under combination with the pronominal stem *ise*, yielding *eese* (← **ey* + *ise*). In Sidaama and Gedeo, the same accusative stem combines with the first-person nominative base (*an-*), yielding *anee* (← **an* + *ee*), with **-ʔ-** epenthesis in Gedeo reflecting the treatment of the lost glide (§3.3.4.1).

In the plural [197], note the preservation of the nasal onset ***n** across the family (see §4.1.1.2). Burji, Haddiya, and Kambaata add *ise* after **ny*, yielding *niis* in Burji; the combination of the singular **ey* and the stem *ise* yields Hadiyya and

Kambaata *nees* (← *ni* + *ise*). Sidaama *ninké* includes the additional clitic **=k**, common in the genitive constructions (§§4.1.4.1, 4.1.4.2), and the **-n-** is likely inserted as a placeholder for the dropped glide **y** (§§3.4.2, 2.1.4). The accusative pronoun of Gedeo *noʔooʔo* shows generalization from the nominative counterpart *noʔo* 'we' with additional suffix **-o** preceded by epenthetic **-ʔ-** (§3.3.4.1) in the accusative. Sidaama has maintained only accent contrast between the nominative and accusative. Below is an overview of the singular and plural accusative and the singular possessive, distinguished only by the exponents preceding ***-y**:

	[103]		[197]		[135]	
PHEC	***e-y**	'me'	***n-y**	'us'	***ʔ-y**	'my'
PNHEC	***ey**	'me'	***n-y**	'us'	***ʔy**	'my'

Second-Person Object

The following list summarizes the second-person accusative forms. The focus is on the inherited accusative stem **ke* and its combination with other pronominal material in language-specific compounds:

	[257]		[196]	
PHEC	***ke**	'you' (sg)	***ki-n-**	'you' (pl)
Burji	**ʃeé**	'you' (sg)	**ʃinsi**	'you' (pl)
PNHEC	***ke**	'you' (sg)	**(*kiʔn-)**	'you' (pl)
Had	**keésẹ**	'you' (sg)	**kiʔn-úwwạ**	'you' (pl)
Kam	**kees**	'you' (sg)	**kiʔneʔéeta**	'you' (pl)
Sid	**ʔaté**	'you' (sg)	**kiʔné**	'you' (pl)
Ged	**ateʔe/atee**	'you' (sg)	**(haʔnooʔo)**	'you' (pl)

Second-person singular accusative forms reflect the same layering seen in the first-person: a stable, inherited accusative stem **ke* plus language-specific combinations (often realized as compounds with other pronominal material). Burji continues *ke* with regular palatalization in the relevant environments (§3.2.1). Hadiyya and Kambaata show vowel coalescence under combination with *ise* (*keese*). Sidaama and Gedeo combine *ke* with the second-person nominative base *ati*, producing compounds that are further shaped by regular lenition and epenthesis processes (§§3.2.2, 3.3.4).

The second-person plural accusative base is segmentable as [196] **ki-n-* with **-n** functioning as the pronominal plurative marker across HEC. The PNHEC cognate **kiʔn-* reflects a generalized second-person plural possessive

pronoun [187] **kin-ʔ-* → **kiʔn-* (with the genitive ***-ʔ**, see §4.1.4.1). The remaining NHEC material is language-specific: Hadiyya adds plurative **-uwwa** and Kambaata adds the accusative **-éeta**. Gedeo added the accusative **-o** to the second-person plural nominative stem [104] *haʔno* 'you' (pl) and generalized it to the accusative (*haʔnooʔo*). These extensions pattern with other plural and case morphology discussed in §§4.1.3, 4.1.5.2, and with the morpheme boundary alternations discussed in §§3.3.2, 3.3.4.1.

Third-Person Object

The third-person accusative series isolates the gender-case vowels most clearly. The list is organized to separate the shared stem **is-* from the accusative vowel exponents and plurative ***-n**.

	[151]		**[292]**		**[86]**	
PHEC	***is-e, *is-é**	'her'	***is-o, *is-ó**	'him'	***is-n-**	'them'
Burji	**ʔiʃeé**	'her'	**ʔisi̱**	'him'	**isinoo**	'them'
PNHEC	***is-e**	'her'	***is-ó**	'him'	***is-n-**	'them'
Had	**ʔíse̱**	'her'	**it'to**	'him'	**(ʔit't'-úwwa̱, it't'u), issʔu**	'them'
Kam	**isé-ta**	'her'	**isú**	'him'	**issoʔóota**	'them'
Sid	**ʔisé**	'her'	**ʔisó**	'him'	**ʔinsá**	'them'
Ged	**isee**	'her'	**isoo, isooʔo**	'him'	**insaʔnee, insaʔné, issaʔné,**	'them'

Third-person accusative pronouns provide the clearest evidence for case-gender vowel exponents. The feminine **is-e* 'her' and masculine **is-o* 'him' share a stem **is-* and contrast only in their terminal vowel, supporting the reconstruction of an accusative gender opposition in PHEC. Hadiyya and Sidaama preserve this opposition most transparently. Kambaata replaces ***-o** with **-ú**, and Burji neutralizes the inherited vowel opposition via analogy with the nominative series, maintaining the accusative-nominative contrast primarily through accentual differences.

The third-person plural accusative form builds on the same stem, with **is-n-* segmentable as 3rd person + plurative ***-n**. Additional suffixes are language-specific and align with the plural extensions already observed in the second-person plural accusative series.

4.1.6.2 Nominative Pronouns

The nominative (subject) pronoun paradigm retains the same person-marking stems as the accusative series, but it replaces accusative case vowels with nominative exponents. The third-person singular is especially informative: *is-* combines with nominative gender vowels (**-i**/**-y** ‘she’, **-u**/**-w** ‘he’) that correspond to the accusative opposition (**-e** vs. **-o**). Plural nominatives consistently retain ***-n** but show extensive generalizations (placed between parentheses), restructuring in stem shape, and addition of suffixes. The following lists of examples present the nominative reflexes with PHEC and PNHEC reconstructions:

First-Person Subject

	[2]		[63]	
PHEC	***an-i**	‘I’	***naʔ-**	‘we’
Burji	**ʔáni̱**	‘I’	**naánu̱**	‘we’
PNHEC	***ani**	‘I’	***naʔ-**	‘we’
Had	**ʔáni̱**	‘I’	**neési̱, neese**	‘we’
Kam	**án**	‘I’	**naʔóot**	‘we’
Sid	**ʔani**	‘I’	**(nínke)**	‘we’
Ged	**ani**	‘I’	**noʔo**	‘we’

The first-person singular nominative **an-i* ‘I’ is uniformly attested across HEC. The stem *an-* is stable, and the final **-i** is consistent with the nominative vowel exponent (cf. the case-vowel discussion in §4.1.5.2).

The plural first-person nominative series supports a common base that is related to the first-person accusative **n-y* (§4.1.6.1), but the surface forms largely reflect combination and analogy. Burji attaches additional pronominal material (*an* ‘I’) to the base, while Sidaama and Gedeo generalize a single stem across the nominative and accusative. Sidaama generalizes the accusative stem to the nominative, maintaining the contrast through accent only: *nínke* (nom) vs. *ninké* (acc). Given these generalizations and restructurings, the plural base is best treated as provisional.[180]

[180] There are additional language pair comparisons: Kambaata and Gedeo share the same terminal vowel **-o** before the addition of **-ti** in Kambaata, so prior to the suffixation; Kambaata *naʔ-* is more conservative. Kambaata **-t** and Sidaama **-k** are referenced to the gender exponents discussed in §4.1.5.1. The **-n-** before the **-k** in Sidaama might be the result of the same process of stem **ʔ** + **k** fusion into **kk** and degemination and epenthesis into **-n-k** that occurred with the genitive in Burji and Gedeo (§§4.1.4.1, 4.1.4.4).

Second-Person Subject

The next list presents the second-person nominative series. The singular is expected to reflect the inherited **at-* base with a nominative vowel, while the plural series is evaluated separately because it shows deeper stem divergence between Burji and the NHEC languages.

	[27]		[104]	
PHEC	***at-i**	‘you’ (sg)	***ati-n-**	‘you’ (pl)
Burji	**aʃi**	‘you’ (sg)	**aʃinu**	‘you’ (pl)
PNHEC	***at-i**	‘you’ (sg)	***aʔ-n-**	‘you’ (pl)
Had	**ʔátị̥**	‘you’ (sg)	**(kiʔnúwị̥)**	‘you’ (pl)
Kam	**át**	‘you’ (sg)	**aʔnoʔóot**	‘you’ (pl)
Sid	**ʔáti**	‘you’ (sg)	**(kiʔne)**	‘you’ (pl)
Ged	**ati**	‘you’ (sg)	**haʔno**	‘you’ (pl)

The second-person singular nominative **at-i* ‘you’ (sg) is securely reconstructed. Burji **ʃ** (← ***t**) reflects regular palatalization before high front vowels (or older ***y**; see §3.1.2.1).

The divergent plural bases **ati-n-* (PHEC) and **aʔ-n-* (PNHEC) between Burji and the NHEC languages suggest either early replacement or restructuring. Kambaata and Gedeo share the provisionally reconstructed **aʔ-n-* with additional suffixes. Burji *aʃinu* ‘you’ (pl) extended with the plurative ***-n** also looks like a generalization of the singular counterpart *aʃi* ‘you’ (sg). But unlike the NHEC stems, it has more EC parallels (see Yaaku, Ts’amakko, Dhaasanac). Therefore, second-person nominative **ati-n-* is provisionally reconstructed for PHEC based on retention in Burji, and **aʔ-n-* for PNHEC based on distinctive cognates in Kambaata and Gedeo. The parallel stem *kiʔn-* in Hadiyya and Sidaama (together with the accusative counterpart [196]) is a generalized variant of the second-person plural possessive pronoun [187] **kin-ʔ-* → **kiʔn-* (with the genitive ***-ʔ**, see §4.1.4.1). The comparison below shows the stem generalization between the NHEC accusative (acc), nominative (nom), and possessive (poss) pronouns:

	[196]	acc	[104]	nom	[187]	poss
Had	**(kiʔn-úwwạ)**	‘you’ (pl)	**(kiʔnúwị̥)**	‘you’ (pl)	**kiʔnúwwi**	‘your’ (pl)
Sid	**(kiʔné)**	‘you’ (pl)	**(kiʔne)**	‘you’ (pl)	**kiʔneʔee, kiʔnee**	‘your’ (pl)
Kam	**(kiʔneʔéeta)**	‘you’ (pl)	**aʔnoʔóot**	‘you’ (pl)	**kiʔneʔée**	‘your’ (pl)
Ged	**(haʔnooʔo)**	‘you’ (pl)	**haʔno**	‘you’ (pl)	**(haʔnnoke, haʔnnot’t’e)**	‘your’ (pl)

The parallels between parentheses reflect generalization (leveling) and paradigm restructuring (reallocation across cells) to fit a different paradigm; for example, Gedeo has reshaped the nominative stem and extended it into the accusative and possessive.

Third-Person Subject

The third-person nominative paradigm tests whether the accusative stem *is-* is retained and whether nominative gender vowels (***-i**/***-u**, with glide variants) are preserved. The plural forms are treated as stem + plurative ***-n** with language-specific extensions.

	[6]		[7]		[11]	
PHEC	***is-i, *is-y**	‘she’	***is-u, *is-w**	‘he’	***is-n-**	‘they’
Burji	**ʔíʃi̱**	‘she’	**(ʔísi̱)**	‘he, it’	**ʔisínu̱**	‘they’
PNHEC	***is-i**	‘she’	***is-u**	‘he’	***is-n-**	‘they’
Had	**ʔísi̱**	‘she’	**(isse/issoʔi, it’t’o)**	‘he/she’ (honorific), ‘he’ (subj)	**(it’t’uw), issuw**	‘they’
Kam	**(íse)**	‘she’	**ís/ísu**	‘he’	**issoʔóot**	‘they’
Sid	**(ise)**	‘she’	**(ʔísi)**	‘he’	**ʔínsa**	‘they’
Ged	**(ise)**	‘she’	**(isi)**	‘he’	**(insaʔnne, insaʔne)**	‘they’

The third-person singular nominative forms preserve the inherited case-gender contrast most transparently: **is-i*, **is-y* ‘she’ (feminine) and **is-u*, **is-w* ‘he’ (masculine). The Burji palatalized reflex **ʃ** in the feminine cognate might be a result of **s** + **y** fusion as proposed in (§4.1.5.2 (subsection on Nominative)). The vowel of the case-gender has been leveled across the feminine and masculine, except in Kambaata. Hadiyya again shows the unique glottalized and geminated **t’t’** that it has in the third-person masculine and plural accusative pronouns, and still retains the older reflexes in the honorific pronoun. The **t’t’** is likely a generalization of the third-person singular (m/f) possessive pronoun stem [279] *ʔít’t’i* (via **is=t-ʔ-*) to all the third-person pronouns except the singular feminine nominative and accusative.

The plural **is-n-* builds on the singular with the plurative ***-n**, with retention in Sidaama, addition of a suffix in Kambaata, and generalization of the possessive stem in Gedeo and Hadiyya (§4.1.4.2). The glottal stop **ʔ** in NHEC, and the individual languages restructured the accusative and nominative counterparts heavily, as the following comparison shows:

	[86]	ACC	[11]	NOM
Burji	**isinoo**	‘them’	**ʔisínụ**	‘they’
Had	**(ʔit’t’-úwwạ, it’t’u, issʔu)**	‘them’	**(it’t’uw), issuw**	‘they’
Kam	**issoʔóota**	‘them’	**issoʔóot**	‘they’
Sid	**ʔinsá**	‘them’	**ʔínsa**	‘they’
Ged	**(insaʔnee, insaʔné, issaʔné)**	‘them’	**(insaʔnne, insaʔne)**	‘they’

The terminal vowels of the plurals are not reconstructed because synchronic exponents diverge across languages. Two candidates emerge from terminal vowel quality: ***-uwa** (back vowels) and ***-eye** (front vowels); e.g., PHEC [11] **isn-uwa* ‘they’ (§4.1.1.2). A broader East Cushitic comparison may provide further evidence for recovering such exponents.

For the reconstruction of the possessive pronouns, see §4.1.4.2.

The personal pronoun paradigms support the following PHEC reconstructions: (a) a shared stem **is-* for the third-person across accusative and nominative; (b) case-marking terminal vowels ***-e** (acc), ***-i** (nom) for the feminine; ***-o** (acc), ***-u** (nom) for the masculine; (c) the plurative ***-n** is segmentable in the plural forms; and (d) the genitive ***-ʔ** deducible from the possessive pronouns. The first and second-person paradigms show greater divergence between Burji and NHEC. There is no evidence of case distinction in the third-person plural forms.

Summary of the strategies applied to the pronouns in HEC.

	Generalization (leveling)	Restructuring (reallocation)	Combination (compounding /fusion)	Addition (extensions)
Burji	x	x	x	x
Sid	x	x	x	x
Ged	x	x	x	x
Had	x	x	x	x
Kam	x	x	x	x

Generalization/leveling and paradigm restructuring are especially visible in the plural series and in the parallels in parentheses. Compounding is exemplified in the discussion on 1sg/2sg. Addition is exemplified by language-specific plural and case extensions.

The preceding sections establish the main nominal markers that can be compared across HEC, including number and nominal extensions, the exponents used in genitive constructions, and the primary gender and case markers. Some patterns recur broadly and support reconstruction to higher

levels, while others are restricted to individual languages or to the PNHEC group and therefore constitute candidates for consideration as later innovations.

The next section turns to verb morphology. The verbal system supplies a second set of reconstructable markers, especially in derivational morphology, and it provides many of the triggers for the morpheme boundary processes described in Chapter 3. Reconstructing verbal suffixes and their allomorphy also helps anchor fossilized morphemes in cognates of the reconstructed lexicon discussed in Chapter 2.

4.2 Verb Morphology

This section reconstructs the major components of HEC verb morphology with an emphasis on derivational markers and support verbs. The analysis compares the formal exponents of passive (§4.2.1), causative (§4.2.2), reflexive (§4.2.3), and denominative (§4.2.5) morphology and evaluates their distribution and conditioning across the five languages, in addition to the support verbs (§4.2.4) used in denominative derivation. There is considerable formal overlap between the denominatives and the other markers. Therefore, some of the examples discussed in one section (e.g., the section on Causative -is) may also (or alternatively) fit in the other (e.g., The Causative Denominative).

Where a marker shows regular allomorphy, the relevant phonological or morphological environments are stated explicitly. Verb morphology is directly tied to the boundary processes documented in Chapter 5, since several derivational suffixes trigger assimilation, palatalization, fusion, and epenthesis patterns at the stem-suffix boundary. For this reason, the discussion keeps the reconstruction of each morpheme aligned with the morphophonological evidence and distinguishes inherited markers from later innovations and co-morphologies (contact-induced).

4.2.1 The Passive

The passive suffix **-am** is attested in Burji (Tesfaye Baye, 2015, p. 181), Hadiyya (Ritter, 2007, p. 6), Kambaata (Treis, 2008, p. 357), Sidaama (Yri & Pepper, 2019; Kawachi, 2007, p. 342), and **-em** is attested in Gedeo

(Wedekind, 2008, p. 2; Gasparini, 1995, p. 97). Therefore, PHEC ***-am** is securely proposed.[181]

	ø/-ʔ		**-am**	
Had	**dun-**	'to leak, to pour'	**dun-am-**	'to be poured, to leak out'
Kam	**dun-**	'spill'	**dun-am-**	'be spilled, be poured'
Sid	**dun-**	'pour out'	**duʔn-am-**	'be poured out'
Ged	**dun-**	'pour out'	**dun-em-**	'flow out, be poured'
Burji	**dunʔ-**	'to poke'	**-**	-

Sidaama *duʔnam-* 'be poured out' (pass), from the derived stem *duʔn-* 'pour out', is identified as autobenefactive in Yri & Pepper (2019). Hudson (1989, p. 89) has Sidaama *dun-am-* 'leak (of e.g. water)' from the underived stem. Sidaama *duʔn-* 'pour out' is a parallel to Burji *dunʔ-* 'to poke'. Alemu Banta's (2016) glossing of the Kambaata passive *dun-am-* 'spilling out, pouring out' is inaccurate and should be 'be spilled/poured out' as in the rest of NHEC.

Gedeo **-e** is also found in the middle variant **-ed'**. Wedekind (1980, p. 149) considers such a correspondence of **a** and **e** to be related to stress placement,[182] e.g., Sidaama *galado* and Gedeo *galedo* 'monkey'.

4.2.2 The Causative

The PHEC causative marker is ***-s**. It is retained as **-s** following single stem-final sonorants, is preceded by an epenthetic **-i-** as **-is** following stem-final geminates and clusters, and in PNHEC, is fused and geminated as ***s-s** following single stem-final obstruents (see §3.1.2.2 for related processes). A palatalized variant **-ʃ** is found in Kambaata, Sidaama, and Gedeo, and an affricated **c** is found in Gedeo.[183]

[181] Data sources: Hadiyya (Ritter, 2007, p. 34), Kambaata (Alemu Banta, 2016; Treis, p.c.), Sidaama (Yri & Pepper, 2019), Gedeo (Wedekind, 2008, p. 48), Burji (Getahun Addo, 2017, p. 224). See also Hudson (1989, p. 89; 2006, p. 296).

[182] "For Sidamo, the basic stress rules can be summarized from Hudson's paper (1969, 13f.): (a) The penultimate syllable of the morpheme is stressed, (b) but some affixes and some verb forms demand word final stress. For Gedeo, the basic rules appear to be the opposite: (a) The final syllable of the word stem is stressed, (b) but most suffixes and some noun and verb forms demand penultimate stress". (Wedekind, 1980, p. 139)

[183] Data sources: Kambaata (Treis, 2022), Gedeo (Eyob Kelemework, 2015), Sidaama (Kawachi, 2007), Hadiyya (Tadesse Sibamo, 2015; Ritter, 2007), Burji (Tesfaye Baye, 2015). Only a few examples are from the database. Where translations are missing, it means they are not provided in the source.

4.2.2.1 Causative *-s

The causative **-s** is less frequently attested than **-is**.

Had	*t'an-* 'do, perform' → *t'an-s-* 'to enable somebody to do something'
	mar- 'to go' → mas-s- 'to take, to let go'
Kam	*zaaz-* 'flow' → *zaas-s-* 'make flow'
	sagab- 'become thirsty' → *sagas-s-* 'make thirsty'
Sid	*makkal-* 'to become crazy/mad' → *makkal-s-*
	mool- 'to become dry' → *mool-s-* / *moo-ʃʃ-*
Burji	*hiik'-* 'close' → *hiisk'-* (with metathesis)
	goob- 'be brave' → *gooss-*, *goobis-*

In Kambaata and Sidaama, stem-final obstruents sometimes fully assimilate to the attached causative **-s** (§3.1.1.3). The full assimilation of labials and alveolars has been preceded by a metathesis process (**C+s** → **sC**) in Burji, which is still retained before velars. Gedeo is the only HEC language that does not have an allomorph **-s** without epenthetic **-i-**.

4.2.2.2 Causative -is

The frequently and widely attested variant **-is** of the causative comprises the epenthetic **-i-** and the causative **-s**.

Had	*holl-* 'chase'→ *holl-is-*
	ʔaanʃ- 'wash' → *ʔaanʃ-is-*
Kam	*birr-* 'stop (of rain)' → *birr-is-* 'take shelter (from rain)'
	c'ung- 'drink the last drop (of a drink)' → *c'ung-is-* 'empty (a container) to the last drop'
Sid	*c'oomm-* 'to taste good' → *c'oomm-i-s-* / *c'oomm-i-ʃ-*
Ged	*fiʃʃ-* 'smell' → *fiʃʃ-i-s-* / *fiʃʃ-i-ʃ-*, *darʃ-* 'swell' (intr) → *darʃ-i-s-* / *darʃ-i-ʃ-* 'swell' (tr)
Burji	*damm-* 'to pass or spend the night' → *dammis-* 'to allow someone to pass or spend night in one's home'

The formal and functional overlap with the derived verb **is* 'to make' also contributes to its prominence across all the HEC languages, as discussed in §3.1.1.3. As already seen in the Sidaama and Gedeo examples above, there is a palatalized variant **ʃ**, which shows free variation with **s**.

4.2.2.3 Causative -ʃ

In Kambaata, Sidaama, and Gedeo, the causative ***s** has been recently palatalized into **ʃ** (§3.1.2.2).

Kam	*ʃiil-* 'melt' → *ʃiiʃ-ʃ-*
Sid	*bub-* 'to burn' → *buʃ-ʃ-*
	duub- 'to become satisfied' → *duuw-i-s-* / *duuw-i-ʃ-* / *duuʃ-ʃ-*
Ged	*utal-* 'jump' → *utal-c-*
	leell- 'appear, be seen' → *leell-i-ʃ-* 'show'

The innovated palatalized **-ʃ** occupies all the environments where the allomorphs of the alveolar **-s** occur, but it is most prevalent after sonorants, particularly **l**, as in Kambaata. The environment after the lateral **l** is also where Gedeo has innovated the affricated variant **c** of the palatalized **ʃ**, e.g., *utal-* 'jump' → *utal-c-*.

4.2.2.4 Double Causative

The double causative (caus1 + caus2) includes a causative and complex form (**-s-y-i-s**) comprising the causative **-s** (caus1) followed by another causative but preceded by a support verb **y** and epenthetic **-i-** yielding **-siis**.

	Stem		**Single causative**	**Double causative**
Had	**ʔag-**	'drink'	→ **ʔag-is-**	**ʔag-isiis-**
Kam	**waal-**	'come'	→ **waaʃ-ʃ- 'bring'**	**waaʃ-ʃ-isiis-**
Sid	**t'orʃ-**	'to pour'	**-**	→ **t'orʃ-isiis- / t'orʃ-iʃiiʃ-**
Ged	**ful-**	'get out'	→ **ful-c-**	**ful-ciis**
	raʔ-	'ripen'	→ **raʔ-is-**	**raʔ-isiis- / raʔ-iʃiiʃ-**
	ib-	'fall'	→ **ib-is-**	**ib-isiis-, ib-is-isiis-**
Burji	**ʔit-**	'eat'	→ **ʔit-is-**	**ʔit-is-iss-**[184]

All the HEC languages have a double causative with the non-palatalized **-siis-**, while multiple causatives of the form **-s-isiis-** are found in Gedeo, Kambaata, and Burji only. Gedeo, Kambaata, and Sidaama have instances of the palatalized **ʃ** (and Gedeo **c**) (caus1) and alveolar **s** (caus2), e.g., **-ʃ-isiis-** in the double/multiple causative formations. Gedeo has a variant **c** that occurs after stem-final **l**, as in *ful-* 'get out' → *ful-c-*, *ful-ciis*. Gedeo and Sidaama have optional alternation between the palatalized harmonized **-ʃiiʃ-** and non-palatalized alveolar **-siis-** allomorphs. Burji and Hadiyya have a

[184] Kambaata also has *it-* 'eat, rip off' → *it-is-* 'feed, offer (food)', *it-siis-* 'cause to be ripped off'.

geminate **ss** in caus2 via the addition and fusion of the verb **y* 'to say, do, be' and the caus2 **-s**, i.e., **-siss**. For example, in Hadiyya, the double causative of *kiʔ-is-* (caus1) 'to make somebody stand up, to wake up' is *kiʔ-isiss-* (caus1 and caus2) 'to wake somebody up' from the stem *kiʔ-* 'to stand up, to wake up'. Otherwise, Hadiyya does not share most of the other double causative formations.

A summary of the allomorphs shared by two or more HEC languages is provided below:

	Single causative	i/y-causative	Double	Multiple
Had, Kam, Sid, Burji	**-s**			
Kam, Sid, Ged	**-ʃ**			
all HEC	**……**	**-is**		
Kam, Sid, Ged	**……**	**-iʃ**		
all HEC	**……**		**-siis**	
Burji, Had	**……**		**-siss**	
Ged, Sid	**……**		**-ʃiiʃ**	
Ged, Kam, Burji	**……..**			**-sisiis**
Ged, Sid, Kam	**……..**			**-ʃisiis**

Multiple causatives are allowed in all HEC languages. However, the palatal variant **-ʃ** of the causative suffix in Kambaata, Sidaama, and Gedeo never occurs as a second causative, i.e., it is not attested as caus2 following and **-s** causative. The reconstructable forms are three allomorphs of the single causative – ***-s**, ***s-s** (PNHEC), and ***-is** – and one double causative ***-siis**. However, some of the other listed double/multiple causative forms might have also existed at the proto-levels.

4.2.3 The Reflexive

The HEC literature variably labels the same exponent as reflexive, middle, or autobenefactive. The use of "reflexive" here as a cover term broadly includes middle and autobenefactive functions (§1.5.2), which are formally and functionally indistinguishable synchronically across HEC languages (§4.2.3.1). In the HEC languages, these functions are expressed by a single morphological marker (**-ʔ** and its allomorphs), making them formally indistinguishable. Glosses in this section occasionally specify the function where it has been provided in the source, but the functional boundaries are fluid, language-specific, and beyond the scope of this study. Formally, the historical proposal assumes that the middle and autobenefactive functions are

served using a reflexive marker ***ʔ** and the combination of ***ʔ** and ***t**, ***k** (***-t-ʔ-**, ***-k-ʔ-** → ***-t't'**, ***-k'k'**).

This section is organized as follows: §4.2.3.1 presents the basic reflexive ***-ʔ**, documenting its distribution, its alternations conditioning (metathesis after sonorants, fusion after obstruents), and the theoretical justification for reconstructing **ʔ** rather than the commonly posited **d'**. §4.2.3.2 examines the complex reflexive **-ak'k'** found primarily in Kambaata and Hadiyya. §4.2.3.3 analyzes the reflexive **-it't'** in Hadiyya and Kambaata as the historical fusion of (agentive) **-t** and reflexive **-ʔ**, presenting synchronic evidence and discussing the limits of stem comparability. §4.2.3.4 addresses the **-Vd'/-d'** variant in Burji, Gedeo, and Sidaama, attributing it to Oromo contact influence and distinguishing it from the inherited PHEC system.

4.2.3.1 Reflexive *-ʔ

The suffix **-ʔ** is found as a reflexive in all HEC languages: Burji (Tesfaye Baye, 2015, p. 183), Hadiyya (Tadesse Sibamo, 2015, p. 131), Kambaata (Treis, 2008, p. 217), Sidaama (Kawachi, 2007, p. 335), and Gedeo (Eyob Kelemework, 2015, p. 190). The productivity and antiquity of this suffix are evident from its systematic participation in morpheme boundary alternation processes: fusion following obstruent-final stems (§3.3.1) and metathesis following sonorant-final stems (§3.3.2). These processes, shared across NHEC languages, indicate that ***-ʔ** has been a stable component of the HEC verbal system since P(N)HEC.

In NHEC, the morpheme boundary alternations of the reflexive suffix **-ʔ** are conditioned by whether it attaches to a stem ending in sonorant or obstruent consonants, i.e., metathesis following stem-final sonorants (§3.3.2) and fusion following stem-final obstruents (§3.3.1).

The following examples illustrate the sonorant-final environment. The presentation contrasts NHEC forms showing metathesis with corresponding stems in Burji, Kambaata, and Gedeo, where metathesis is absent:[185]

[185] Data sources (additional data): Gedeo (Wedekind, 2008, p. 19), Kambaata (Alemu Banta, 2016), Burji (Getahun Addo, 2017, p. 306).

			[36]				[120]	
Burji	**gam-**	‘to bite’	**-**	-	**fil-**	‘to comb’	**silʔ-**	‘to comb for oneself’
Had	**-**	**-**	**gaʔm-**	‘to bite’	**-**	**-**	**-**	-
Kam	**gam-**	‘cause bellyache’	**gaʔm-**	‘to bite’	**-**	**-**	**fiil-**	‘unravel’
Sid	**-**	**-**	**gaʔm-**	‘to bite’	**-**	**-**	**-**	-
Ged	**-**	**-**	**gaʔm-**	‘to bite’	**fil-**	‘comb’	**fiʔl**	‘to comb’

The sets show that metathesis is characteristic of the NHEC reflexive system, whereas the corresponding cognate [120] in Burji does not show metathesis. Where no underived stem is attested, the most economical interpretation is fossilization rather than productive derivation.[186]

When **-ʔ** follows a stem-final obstruent, glottalization occurs through fusion (§3.3.1). This process results in the root-final consonant becoming glottalized and geminated (**C’C’**).

b, d/t, k/g, ʃ/j/z/s + ʔ → p’p’, t’t’, k’k’, c’c’

The following examples give representative stem-reflexive pairs showing the regular fusion outcome (**C + ʔ → C’C’**) across NHEC:[187]

	Stem		**Reflexive**	
Had	**amad-**	‘to catch’	**amat’t’-**	‘to hold some parts of the body’
Kam	**kuʃu**	‘choke’	**kuc’c’u**	‘choke, cough’
Sid	**daak-**	‘to grind’	**daak’k’-**	‘to grind (for oneself)’
Ged	**dib-**	‘prepare ensete food’	**dip’p’-**	‘prepare food for oneself’
cf. Burji	**ʃaf-**	‘carve’	**ʃad’d’-**	‘to carve for oneself’

As shown above, Burji has a unique development **f+d’, b+d’** → **d’d**’ as seen in *ʃad’d’*- ‘carve for oneself’ with exceptions, e.g., *gaaf*- ‘grind’ → *aafd’* ‘grind

[186] Burji shows cases of free variation between the allomorphs **ʔ** and **d’**, and Sasse (1982, p. 18) notes that Burji **d’** and **ʔ** are dropped before consonants with oncomitant lengthening of the preceding vowel.
[187] Data sources: Gedeo (Wedekind, 2008, p. 58), Sidaama (Kawachi, 2007, pp. 57–58, 35; Indrias Xa’miso et al., 2007, p. 101), Hadiyya (Tadesse Sibamo, 2015, p. 133; Ritter, 2007, p. 7), and Kambaata (Treis, 2008, pp. 243, 269, 242, 267, Treis, p.c.), Burji (Getahun Addo, 2017, p. 777).

for oneself' (Tesfaye Baye, 2015, pp. 182–183) and shows no evidence of a fusion process similar to that found in the NHEC languages.

This analysis differs from all the existing HEC literature (at least on Gedeo and Sidaama) in proposing the glottal stop **ʔ** as the sole variant responsible for fusion and glottalization instead of the commonly posited **d'**. Five arguments converge to support reconstructing ***-ʔ**:

(a) Parsimony: There is no reason for postulating **d'** with subsequent change to **ʔ**, then metathesis and fusion, when an established allomorph **ʔ** already exists in the paradigm. Occam's Razor favors projecting the attested **ʔ** directly to the proto-level rather than positing an intermediate ***d'**.

(b) Traces of assimilation: The hypothetical allomorph ***d'** lacks the expected traces of partial assimilation it should have left on adjacent **m**, in the form of **nd'**, before changing to **ʔ**, that other alveolars have left (§3.1.1.4). Comparative evidence from NHEC shows consistent alveolar assimilation (e.g., ***-m** → **-nd-** / __ **d**); the absence of **nd'** patterns undermines the ***d'** hypothesis.

(c) Phonetic plausibility: The fusion of the voiceless glottal stop **ʔ** with both voiced and voiceless obstruents is phonetically more plausible in terms of yielding regular voiceless glottalized geminates **C'C'** than fusion with the voiced implosive **d'** (§3.3.1). Therefore, it is a fusion of **ʔr** that yields the expected glottalized **d'(d')** in Gedeo and Sidaama (§3.3.1.5).

(d) Phonemic status: There is limited evidence to establish **d'** as a stable phoneme in minimal pairs in the relevant HEC languages (Tesfaye Baye, 2015, p. 27; Kawachi, 2007, p. 35). If **d'** were a core proto-phoneme, one would expect robust paradigmatic contrast; instead, it shows restricted distribution and contextual variation.

(e) Distribution: The consonant **d'** is restricted to Burji, Sidaama, and Gedeo, i.e., it is not reconstructable to PHEC or PNHEC. Its presence is attributed to alternation process (**ʔr** → **d'**) documented in §3.3.1.5 and Oromo contact influence (§4.2.3.4). This geographically/historically restricted pattern argues against a PHEC origin.

However, suppose there is such an alternation **d' ~ ʔ**. In that case, it must be confined to the environment following a sonorant and must have been

innovated individually due to the foreign nature of the borrowed sound **d’**. Synchronically, **-ʔ** is formally confused with the variant **-d’** because they are functionally indistinguishable (middle ~ reflexive ~ autobenefactive, where those terms are used equally to refer to form and function).

4.2.3.2 Reflexive -ak’k’

The reflexive **-ak’k’** is found in Kambaata (Treis, 2022, p. 17; Treis, 2008, p. 217; Hudson, 1976, p. 272) and Hadiyya (Tadesse Sibamo, 2015, p. 134; Sim, 1991, pp. 54–55, fn. 6). Following is a representative list of Kambaata and Hadiyya examples, organized by stem and the derived reflexive form:[188]

	Stem			**Reflexive**
Kam	**eeb-**	‘to bring’	→	**eeb-ak’k’-**
	maat’-	‘to hide’	→	**maat’-ak’k’-**
Had	**ʔeeb-**	‘to bring/marry’	→	**ʔeeb-ak’k’-**
	mooʔ-	‘see’	→	**mooʔ-ak’k’-**

The distribution of **-ak’k’** is currently limited to Kambaata and Hadiyya. Based on the present evidence, this pattern is best treated as a shared innovation (see §4.2.5.5) rather than a proto-level reflexive exponent.

Interestingly, older literature does not report Hadiyya **-ak’k’**. Plazikowsky-Brauner (1957, p. 29) records a reflexive **-ak** in Hadiyya, e.g., *uw-ak* ‘to give for oneself’. Hayward (1984, p. 90) considers those Kambaata and Hadiyya forms with **-ak’k’** to be innovations. The older description of **-ak** for Hadiyya may reflect underdocumentation, variation, or different segmentation in the earlier sources. However, the question would remain: Is Hadiyya **-ak’k’** a recently introduced loan via contact with Kambaata, or simply unrecorded in the previous studies? For the morphophonological behavior of the reflexive **ʔ** in fusion processes with obstruents and metathesis with sonorants, see §§3.3.1, 3.3.2.

4.2.3.3 Reflexive -it’t’

Hadiyya and Kambaata show synchronic evidence of **-it’t’** used in middle forms of words ending in consonants other than **t**. The pattern can be analyzed as the historical fusion of (agentive) **-t** and reflexive **-ʔ**, yielding a glottalized geminate outcome. The sources for these languages do not recognize this

[188] Data sources: Kambaata (Treis, 2022, p. 17; Treis, 2008, p. 193) and Hadiyya (Tadesse Sibamo, 2015, p. 134).

ending as a suffix, so this section presents key examples that motivate segmenting **-it't'** as a complex reflexive ending.

The following examples provide comparisons between reflexive forms in **-it't'** and their most plausible underived stems, illustrating both the complexity of the reflexive ending and the limits of stem comparability in some items:[189]

	Reflexive -it't'		**Stem**	
Had	**uull-it't'-**	'to erect, to put an end' (tr)	**uull-**	'to stand (up), to cease, to stop, to take place (intr)'
			uull-is-	'to stop (tr), to make cease, to make stop (intr)'
Kam	**uurr-it't'-**	'keep still/away'	**uurr-**	'stand up, take place'
			uurr-is-	'stand, erect, stop'

This example shows that **-it't'** occupies the same suffixal slot on the stem as the causative counterpart **-is**.

	Reflexive -it't'		**Stem**	
Had	**k'iʃ-it'-**	'to save, to be kind, to grant forgiveness' (intr/tr)	**k'iiʃ-**	'to make a fence around something, to enclose' (intr/tr)
Kam	**k'iʃ-it't'-**	'have mercy'	**-**	-

While Kambaata shows the regular reflex **-it't'**, Hadiyya has a degeminated **-it'** in this example. The same applies to the following example:

	Reflexive -it't'		**Stem**	
Had	**fir-iit'-**	'to gallop' (intr)	**fir-**	'to go out, to exit, to come out, to get out, to climb, to ascend, to pass' (intr)
Kam	**fir-iit'-**	'gallop (of a horse)'	**firr y-**	'jump in different directions'
			firiic'c'a	'gallop' (m)

This example exhibits that the lexical comparability between derived and underived stems varies across items. The vowels on the reflexive can be interpreted as **y** + **-it'** → **-iit'** as proven by the corresponding Kambaata compound verb stem *firr y-* 'jump in different directions' (see §4.2.4.1). The

[189] The data used in this section comes from Hudson (1989), Ritter (2007), Tadesse Sibamo (2015), Treis (p.c.), and Treis (2008).

same applies to the corresponding derived Kambaata *firiic'c'a* 'gallop' (← **y + -is + ʔ**).

As most examples show, this complex variant of the reflexive is usually preceded by epenthesis **-i-t't'** (§3.4.1). Although the first example is straightforward, not every pair securely reflects a transparent derivational relationship, so individual etymologies should be treated cautiously. In other words, some of the above stems might show distant or no semantic relation to the corresponding reflexive marked stem, e.g., the reflexive stems Hadiyya *fir-ii-t'-* (intr) 'to gallop', Kambaata *fir-ii-t'-* (intr) 'gallop (of a horse)' and the corresponding Hadiyya *fir-* (intr) 'to go out, to exit, to come out, to get out, to climb, to ascend, to pass', Kambaata *firr y-* 'jump in different directions'.

Such a form of reflexive marking has limited attestations in Kambaata, and most of the evidence comes from Hadiyya. Below are further examples from Hadiyya:

Reflexive -it'(t')		**Stem**	
haabb-it't'-	'to comb one's hair'	**haabbis-**	'to comb'
c'aakk-it't'-	'to light oneself'	**c'aakk-**	'to light'
ʔap'p'-it't-	'hit oneself'	**ʔap'p'is-**	'to hit'
huuʃ-it't'-	'miss'	**huuʃ-**	'miss, make mistake'
iibb-it't'-	'to get warm'	**iibb-i-s-**	'to heat, to warm (up)'
		iibb-	'to be hot'

Taken together, the Hadiyya data suggest that **-it'(t')** is at least partly productive and interacts with the same morpheme boundary alternations pressure discussed for ***ʔ** (e.g., avoidance of marked consonant clusters via epenthesis and fusion) (see §4.2.5.5).[190]

In *ʔap'p'-it't'-* (refl) 'hit oneself', in addition to the fused Hadiyya **t't'**, multiple fusion processes are observed in the cognate set provided in Hudson (1989, p. 43) (sometimes on the same stem):

Sid		**kibb-** 'hit with horns', **kibb-am-** 'hit with horns (of rams)'
Had	**bb + ʔ → p'p'**	**ap'p'-is-** 'collide (with)'
	t + ʔ → t't'	**ap'p'it't'-**, **ap'p'it't'-am-** 'collide (with)'
Kam	**s/ʃ + ʔ → c'c'**	**k'oc'c'-**, **k'oc'c'-ak'k'-am-** 'collide (with)'

[190] Th stems with **-it't'** are not consistently intransitive, as one would expect in the reflexive derivation, e.g., *sas-s-iit't'-* 'to make something cooked', *sar-* 'to cook, to bake', and *sigg-it't'-* 'to cool down' (tr), *sigg-* 'to be(come) cool/cold' (intr). Some of these derivations might involve a different process, e.g., a reflexive denominative (§4.2.5.5) in the second example, where the semantics are not always transparent.

The table indicates that glottalization outcomes in this set are not limited to a single consonant class and may involve multiple morphophonological pathways (including fusion and reanalysis), sometimes within the same lexical stem. This variability motivates the discussion below of competing segmentations and historical sources.

Tadesse Sibamo (2015) does not provide any explanation except that "in the process of middle (autobenefactive) stem formation the root final obstruents undergo glottalization. i.e., … *t'* to *tt*'", although the above roots do not end in the claimed **t'** (except another example *t'oreet'- '*compete' → *t'oreet't'-* (refl)).

Kambaata also has *hujat't'-* (mid) 'work for one's own benefit', *hujat-* (tr) 'work', and *huj-íta* (f.acc) 'work', which is comparable with Oromo *hòjjed'd'-à* 'work, make', *hòjjéd'd'a* 'I work' (Lloret, 1988, p. 188). However, the question is whether Kambaata **-t't'** has been introduced recently via contact with Hadiyya or has simply not been recorded in previous studies.

4.2.3.4 Reflexive -Vd', -d'

Parallel to the unique reflexives **-ak'k'** / **-it't'** in Hadiyya and Kambaata, Sidaama, Gedeo, and Burji have a reflexive variant with the consonant **-d'**. Burji allomorphs of this variant include **-ad'** and **-d'** (Tesfaye Baye, 2015, pp. 181–182), Gedeo has **-ed'** (Wedekind, 2008, p. 2; Eyob Kelemework, 2015, p. 190), and Sidaama has **-i-d'** (Kawachi, 2007, p. 335; Anbessa Teferra, 2014, p. 50; Yri & Pepper, 2019). The following are the main allomorphs (**-d'**, **-ad'**, **-ed'**, **-i-d'**) illustrated with representative stems and derived reflexive forms in Burji, Gedeo, and Sidaama:[191]

		Stem			**Reflexive**
Burji	**-d'**	**gaaf-**	'grind'	→	**gaafd'**
	-ad'	**daah-**	'swim'	→	**daah-ad'-**
Gedeo	**-ed'**	**erg-**	'send'	→	**erg-ed'-**
Sidaama	**-i-d'**	**abb-**	'to bring'	→	**abb-i-d'-**

The prominence of the variant **-Vd'** in descriptive grammars of Burji and Gedeo (Eyob Kelemework, 2015, p. 71; Tesfaye Baye, 2015, p. 183) may reflect the fact that this variant can attach to stems without triggering epenthesis, fusion, or metathesis, making it morphotactically simpler than the NHEC pattern. In Oromo, the first-person middle is marked by **-ad'd'** (Ali &

[191] Data sources: Burji (Tesfaye Baye, 2015, p. 182), Gedeo (Eyob Kelemework, 2015, p. 192), and Sidaama (Kawachi, 2007, p. 42).

Zaborski, 1990, p. 22), **-ed'd'** (Orma) (Stroomer, 1987, p. 153), and **-d'd'** (Lloret, 1988, p. 189). Gemination aside, the Oromo variants appear to mirror those of Burji **-ad'** and **-d'** and Gedeo **-ed'** in both consonants and vowels.[192] Despite the existence of the Oromo variant with mid-vowel **e**, it is not uncommon for Gedeo to raise **a** to **e** in suffixes; for instance, the passive ***-am** became Gedeo **-em** (§§2.2.1.1, 4.2.1). That Sidaama **-i-d'** is preceded by epenthetic **-i-** to avoid a **CCC** sequence when added to a stem reveals the underlying quantity of **d'** (← **ʔr**). As argued in §4.2.3.1, the variant **-d'** is formally and historically different from **-ʔ**.

Sidaama and Gedeo additionally suggest an alternation between **ʔr** and a glottalized **d'** (often with unstable quantity **d'd' ~ d'**) in the reflexive. In line with §3.3.1.5, this pattern can be interpreted as a comparatively recent, contact-induced reanalysis of the sequence **ʔ+r** (expected under sonorant + **ʔ** metathesis) as a glottalized **d'(d')**, which also helps explain why quantity is least stable here. Unlike the other non-pulmonic consonants, the borrowed **-d'** is hardly found geminated. Additionally, Plazikowsky-Brauner (1957, p. 29) argues that the reflexive **d'** in Sidaama results from Oromo influence, a hypothesis that aligns with the linguistic history and context of Oromo-Sidaama contact in the region. Therefore, Burji **-ad'** and **-d'** are entirely from Oromo, while the Gedeo and Sidaama **-Vd'** variants are from a morpheme boundary alternation process induced by contact with Oromo (§3.3.1.5).[193]

The Burji forms are argued to be direct loans because Burji does not participate in the fusion process responsible for introducing **d'(d')** in the NHEC languages. Burji still shows cases of free variation between the borrowed reflexive **-d'** and the retained **-ʔ**, e.g., *kanʔi*, *kand'i* 'here it is!'; *sil-*, *silʔ-* / *sild'-* 'to comb'; and *kurʔ-*, *kar-d'-* 'decide'.[194]

Lloret's (1988) analysis of Oromo provides complementary evidence for understanding the middle system in EC. She proposes that the seemingly

[192] The degemination itself could be proof of the borrowed origin of the non-pulmonic **d'** in HEC, since all other non-pulmonic consonants can occur as geminated and in the **nd'** cluster. Likewise, if one accepts the **d'** variant for Sidaama and Gedeo, its purported change into a glottal stop is further proof of its external origin.

[193] According to Wedekind (1980, p. 134), Cerulli equates the Gedeo implosive **ɗ** with the geminated glottalized **t't'** (1929, p. 235). However, Wedekind argues that it is only true on the morpho-phonemic level.

[194] Such free variation between **d'** and **ʔ** following alveolar sonorants is not confined to verbs in Burji, e.g., *k'anʔ-i*, *k'and'-i* 'clitoris'. A similar **ʔ** and **d'** alternation before alveolars is found in the South Omotic languages, e.g., Hamar *guldánti ~ gulʔánti* 'belly button', Ari *gulʔa* 'navel' (Petrollino, 2016, pp. 16, 274).

geminated **d'd'** reflexive marker in Oromo derives from an underlying /**ʔt**/ sequence, where the /**ʔ**/ has no underlying weight. She argues that it explains the morphophonological alternation in Oromo between **d'd' ~ t** without requiring compensatory lengthening: "[W]eightless /ʔ/ is only syllabified when it merges with a following consonant, resulting in [d'd']. If it does not merge, it simply deletes without compensatory lengthening" (Lloret, 1988, p. 198). This analysis, though admittedly abstract, provides an explanation for the synchronic patterns observed in Oromo.

In the broader East Cushitic context, there has been debate regarding the historical development of these middle/reflexive markers. Hayward (1984) proposes that the Proto-East Cushitic system alternated between ***ɗ (= *d')** and ***t** in its middle voice paradigm. This alternation underwent different patterns of leveling in various branches: "[I]n the Highland Eastern Cushitic languages... levelling also took place in the suffixed middle root extension, but in favour of ***ɗ**" (Hayward, 1984, p. 84). This contrasts with the PNEC branch, which leveled toward ***t** instead.

4.2.3.5 Summary and Reconstruction

The comparative evidence presented in §§4.2.3.1-4.2.3.4 supports the following reconstruction of the HEC reflexive system:

> Core PHEC marker: ***-ʔ** is reconstructable to Proto-HEC based on its HEC-wide distribution (attested in all five languages), regular NHEC morpheme boundary alternation (metathesis following sonorant-final stems, fusion following obstruent-final stems), phonological plausibility (voiceless **ʔ** fusing with voiced/voiceless obstruents → voiceless glottalized geminates **C'C'**), and absence of the traces of assimilation (***md'** → ***nd'**) that would be expected if **d'** was the allomorph in the metathesis (§3.1.1.4).
>
> PNHEC innovations: metathesis of sonorant + **ʔ** → **ʔC**, fusion of obstruent + **ʔ** → **C'C'**, and the complex forms **-ak'k'/-it't'** limited to Kambaata/Hadiyya; plausibly analyzable as **-Vk + -ʔ** → **-Vk'k'** / **-Vt + -ʔ** → **-Vt't'** with subsequent reanalysis.
>
> Secondary developments: **-Vd'/-d'** variants in Burji, Gedeo, and Sidaama are attributed to direct borrowing from Oromo (Burji **-ad'/-d'**). There has been contact-induced reanalysis of the sequence **ʔr** as **d'** (Gedeo/Sidaama **-ed'/-i-d'**), hence it is always preceded by a vowel.

> Thus, the phonemic status of **d’** is that it is not reconstructible to PHEC or PNHEC, but limited to individual languages.
>
> Theoretical innovation: This analysis differs from existing HEC literature (Hudson, 1976; Kawachi, 2007; Kelemework, 2015) in proposing **ʔ** as the sole PHEC variant responsible for fusion/glottalization, rather than alternating **d’ ~ ʔ**. The evidence strongly favors projecting ***ʔ** to PHEC, with **d’** emerging later through contact and morphophonological reanalysis.

The next sections (§§4.2.4–4.2.5) examine the support verbs and denominative derivation, where the verbal suffixes (§§4.2.1–4.2.3) participate in the derivation process as part of the denominative suffix attached to nominal stems.

4.2.4 Support Verbs

This section examines two PHEC support verbs: [425] **ʔ* ‘to do’ and [388] **y* ‘to say, do, be’. They differ fundamentally in their inherent transitivity and combine with nominal stems to form compound predicates. Unlike the grammaticalized verbalizer ***ʔ** examined in §4.2.5, these verbs retain some autonomous lexical capacity while participating in derivational structures (see Hudson, 1976, p. 273; Crass, 2007, p. 30; Treis, 2024, p. 165). The transitivity distinction (§4.2.4.3) determines which causative suffix (**ass-** vs. **is-**) a derived stem selects, a pattern that extends through the valency-changing morphology described in §§4.2.1–4.2.3. First, §4.2.4.1 presents the intransitive support verb **y*; §4.2.4.2 then presents the transitive support verb **ʔ*; then, §4.2.4.3 demonstrates their stem-combining patterns across HEC languages.

PHEC [425] **ʔ* ‘to do’, with its derived causative **ass-* (← **ʔs-*), and PHEC [388] **y* ‘to say, do, be’, with its derived causative **is-* (← **y-s*), appear in all HEC languages as independent verbs.

Following Treis (2024), the term “support verb” is used here as equivalent to the term “light verb” of the wider literature. The term “verbalizer” is reserved for the grammaticalized reflex of support verbs when functioning as derivational morphemes on nominal stems. Support verbs in PHEC function simultaneously as independent lexical verbs and as grammatical verbalizers, enabling category change from nominal to verbal stems. This dual function establishes the foundation for the grammaticalization pathway documented in §4.2.5, where support verb constructions evolve from simple denominatives

(§§4.2.5.1–4.2.5.2) through layered combinations (§§4.2.5.3–4.2.5.7) to generalized valency-changing morphology (§§4.2.5.8–4.2.5.9).

The following subsections present each support verb first as an independent lexical item (with cognates demonstrating Proto-East Cushitic inheritance), then examine its grammaticalized function as a verbalizer in denominative derivation. Understanding these dual capacities is essential for interpreting the reanalysis framework in §4.2.5.9, where traditionally analyzed "epenthetic" vowels (**-i-**, **-a-**) are shown to be grammatical reflexes of support verbs **y* and **ʔ*, respectively.

4.2.4.1 The Verb **y* 'to say, do, be'

The PHEC verb [388] **y* 'to say, do, be' is found in all the HEC languages and has a derived causative counterpart **is-* (← **ys-*) 'to make'. The following set presents the stem and causative forms of **y* across the HEC languages. The causative variants show the morpheme boundary alternant after adding the causative suffix ***-s** to the root **y*.

	Stem		**Causative**	
PHEC	***y**	'to say, do, be'	***ys-** → ***is-**	'to say, make'
Burji	**iy-**, **i-**	'say, do'	**iss-**	'do, make'
PHEC	***y**	'to say'	***ys-** → ***is-**	'to say, make'
Had	**y-**	'say, be'	**iss-**	'do'
Kam	**y-**	'say'	**y-isiis-**	'to make say'
Sid	**y-**	'say, judge'	-	-
Ged	**hiyy-**	'say, try, do'	-	-

Several regular patterns emerge from the data. The Burji root is transcribed as *iy-* and *i-* with the causative *iss-*, with phonetic epenthesis of **-i-** in the stem, while the NHEC languages consistently show *y-* as the stem reflex. Gedeo *hiyy-* has inserted initial **hi**; therefore, Eyob Kelemework (2015, pp. 206, 82) observes that, in support verb constructions following nominal stems, the initial **hi-** is deleted (*hiyy-* → *-yy-*).[195] Kambaata **y-isiis-** 'make say' represents a morphological analysis of root + double causative (§3.1.2.3), contrasting with the causative forms **iss-** found in Burji and Hadiyya. This reflects an innovation in Kambaata and fusion in Burji and Hadiyya (**y + s** → geminate **ss**). Sidaama and Gedeo lack attested independent causative forms of **y*, though its use as the support verb persists (§4.2.4.3).

[195] A similar insertion process is found in Sidaama *hiyy-essa* 'wolf, jackal' from PHEC **yeyy-* 'wolf, jackal' (see §2.1.5.3 for initial **h** insertion).

The intransitive patterning of **y* has structural consequences throughout the denominative system. In §4.2.5.2, **y* serves as the foundation for denominal verbs with intransitive or middle semantics (inchoative, stative). The reflex **i-** appearing in causative denominatives like Sidaama *araar-(i)-s-* 'mediate, reconcile' (§4.2.5.7) preserves the support verb as a verbalizer ***y**, contrasting with the causative **-is** suffix on non-derived verbs where **-i-** is epenthetic (§3.1.2.2). This distinction enables denominative origin to be detected even when semantic compositionality has been obscured through lexicalization (§4.2.5.9 discusses semantic specialization).

The generalization pathway documented in §§4.2.5.6-4.2.5.84.2.5.7 argues that denominative constructions built on ***y** + causative **-s** (***y+s**) evolved into the synchronic causative morpheme **-is** through functional expansion, explaining why the causative **-is** is the most prominent, and it formally resembles the denominative pattern (i.e., due to their shared origin and convergence).

The HEC support verb **y* 'to say, do, be' has cognates throughout East Cushitic. Sasse (1982, p. 102) reconstructs PEC **-iy(y)-* 'say' based on widespread reflexes: Burji *iy-*, Sidaama *i-*, *iy-*, Gedeo, Kambaata, Hadiyya, *i-* 'say', Saho-Afar *y-* (suffixing), Afar *-iy(y)-* (prefixing), and possibly Dullay *iy-anna* 'inquit, that is, I mean'. Evidence for this verb persists in fossilized forms across Eastern Cushitic, such as in the verb [21] meaning 'to listen, hear'. The use of **y* as a verbalizer has also been reported for Southern Cushitic (Ehret, 1987, p. 9).

4.2.4.2 The Verb *ʔ 'to do'

The independent PHEC verb [425] **ʔ, *ass-* (caus) 'to do' is directly attested in Hadiyya and Kambaata. Burji, Sidaama, and Gedeo do not preserve **ʔ* as an independent lexical verb, though fossilized traces appear in derivational morphology (§4.2.5.1). Additional HEC cognates strengthen the reconstruction: K'abeena (Crass, 2005, p. 226), Alaaba (Schneider-Blum, 2007, pp. 277–278; 304–305), and Libido (Crass, n.d., p. 30) all attest reflexes of **ʔ* 'to do' (as cited in Treis, 2024, p. 165).

The following set presents the PHEC and PNHEC reconstructions for the verb **ʔ* 'to do' alongside attested reflexes in Hadiyya and Kambaata. The comparative evidence suggests that **ʔ* was the earliest form, with vocalic support developing independently in the daughter languages to satisfy phonotactic needs.

	Stem		**Causative**	
PHEC	***ʔ**	‘to do’	***ʔ-s-** → ***ass-**	‘to do’
PNHEC	***ʔ**	‘to do’	***ʔ-s-** → ***ass-**	‘to do’
Had	**aaʔ-**	‘do’	-	-
Kam	**aʔ-**	‘do’	**ass-, ees-**	‘do’
Sid	**-**	-	**ass-**	‘do’
Ged	**-**	-	**(h)ass-**	‘do’
Alaaba	**aʔ**	‘do’	**ass-**	‘do’
K’abeena	**aʔ-**	‘do’	**ass-**	‘do’
Libido	**aaʔ-**	‘receive’[196]	**-**	-

Allomorphic forms (*ʔ, aʔ, ʔa, a*) reflect variation in the position and realization of the glottal stop relative to the supporting vowel. The causative variants show **ass-*, derived from **ʔ-s-* via fusion and gemination with the causative ***s** (see §3.1.1). Gedeo *hass-* ‘do’ (parallel to *hiyy-* ‘say’) has initial **h** insertion. The distribution gap in Burji, Sidaama, and Gedeo reflects lexical replacement rather than morphological loss, as denominative uses persist (§4.2.5.1).

The restriction of **ʔ* to transitive contexts (contrasting with intransitive **y*, §4.2.4.1) is semantically coherent: ‘do’ inherently requires a patient argument, while ‘say’ and ‘be’ pattern intransitively in HEC. This transitivity distinction carries over to support verb contexts (§4.2.4.3) and denominative derivations (§4.2.5.1), where **ʔ* licenses transitive morphology.

4.2.4.3 Stem + Support Verb

As mentioned above, the two verbs can be used as support verbs when they occur in compounds following other stems (or attested nominal stems). They are found separately or directly attached to the preceding stem, where some morpheme boundary alternation processes may occur.

[196] Concerning the meaning of Libido *aaʔ-* ‘accept, receive’, Treis (2024, p. 165, fn. 24) rejects Crass’s (n.d., p. 30) claim that such a support verb used in transitive contexts in Libido goes back to ‘accept, receive’ in light of the comparative data and proposes that it might historically originate from the verb ‘to do’. Of course, Libido *aaʔ-* ‘accept, receive’ could be considered as a semantic extension, as evident in the derived PNHEC verb [32] **ʔ-ʔ-* ‘to take, receive’ → Hadiyya *haʔ-*, *aaʔ-* ‘receive, take, accept, buy’, Kambaata *aak’k’-* ‘receive, take, accept’, Sidaama *haad’-*, *aad’-* ‘take, receive’, and Gedeo *ad’d’-* ‘receive’, with optional initial **h** epenthesis and suffixation of one of the reflexive suffixes (§4.2.3). The stem of the reconstructed form **ʔ* resembles that of the verb **ʔ* ‘to do’ as do its reflexes. Most likely, it has received the reflexive ***-ʔ** in the derivation process, as evidenced by the double vowels and the fused consonant of the attached reflexive suffixes.

The lack of Burji, Sidaama, and Gedeo cognates in the previous section could be amended by provisionally inferring variants which occur as supporting verbs from some verb derivations, at least in Burji and possibly Sidaama, e.g., in Burji:

> *d'agal-* 'to return' → *d'agal-ʔ-* 'to make one's wife return after having left her',
>
> *mur-* 'to cut, harvest' → *mur-ʔ-* 'to delimit, define', *mur-am-* (pass) 'to be cut, harvested', *anga murʔ-* (phrase) 'to kill' (*anga* 'hand'),
>
> *koraa* (m) 'saddle' → *kor-ʔ-* 'to saddle',
>
> *daaba* (f) 'horizon' → *daaban-ʔ-* 'to go horizontally', *daabanee* (adj) 'horizontal', *daaban-s-* (caus) 'to lay horizontally, down' (Roba Dame & Wedekind, 2008).

The Burji reflexes in the examples prompt the inclusion of the variant *ʔ(a)-* as one of the realizations of PHEC **ʔ* 'to do' above.[197] Sidaama has even more tentative evidence for a potential parallel supporting verb *aa-* 'to make, take, give', e.g.:

> *ʃeemaate* 'agreement, pact' → *ʃeemaate aa* 'to make/conduct an agreement with a contractor', *ʃeemaata* 'to make an agreement/contract',
>
> *k'oricca* 'medicine, measure or solution' → *k'oricca aa* 'to take measure, to punish',
>
> *biddiʃʃa* 'guide, directive, department' → *bidiʃʃa aa* 'to give directives, instruction' (Indrias Xa'miso et al., 2007).

The dictionary definition of Sidaama *aa* is 'to give' only (Indrias Xa'miso et al., 2007, p. 60), but the meaning provided in the above examples suggests extending the definition to include 'to make' and 'to take'.

---

[197] The identical form of the reflexive ***-ʔ** and the verb **ʔ* 'to do' makes it difficult to distinguish them when the latter is used as a support verb. However, at least the semantics of some of the provided transitive verb examples do not fit the reflexive/middle semantics, so one can link them back to the verb **ʔ* 'to do'. The support verb also differs in the variants with a vowel, which are not typical of ***-ʔ**.

Semantically, the verb *y-* is intransitive while *aʔ-* is transitive in Kambaata (Treis, 2008, p. 87); *hiyy-* 'do' is intransitive while *ass-* 'say, do' is transitive in Gedeo (when they are part of derived/verbalized verbs) (Eyob Kelemework, 2015, p. 206); and *y-* 'to say' is intransitive while *ʔáss-* 'to do' is transitive in Sidaama (Yri & Pepper, 2019). They are summarized below.

	Intransitive		**Transitive**	
Burji	**iy-**	'say, be'	**iss-**	'do, make'
Kam	**y-**	'to say'	**aʔ-, ass-**	'to do'
Had	**y-**	'to say'	**aaʔ-, iss-**	'do'
Sid	**y-**	'say, try, do'	**ass-**	'do'
Ged	**hiyy-**	'to say, do'	**(h)ass-**	'do'

Accordingly, the root **y* 'to say, do, be' is inherently intransitive, while **ʔ* 'to do' is inherently transitive. The inherited semantics determine each verb's selection of valency suffix (e.g., **ʔ** + passive/causative and **y** + causative) when derived, so they have historical significance, which will be addressed in the following section (4.2.5).

Below is a summary of the two verb roots, their causative stems, and their common reflexes and realizations as independent and cliticized support verbs when attached to another stem:

Root		**Reflexes**	**Causative Stem**		**Reflexes**
***y**	'to say, do, be'	**y, i, iy**	***ys-** → ***is-**	'to make'	**iss-, is-**
***ʔ**	'to do'	**ʔ, a, aʔ, ʔa**	***ʔ-s-** → ***ass-**	'to do'	**ass**

A number of comparative sets exhibit such compounds (or verbalized stems) in the individual HEC languages. The examples show that, in compound forms with both **y* 'to say, do, be' and **ʔ* 'to do' functioning as verbalizers are found, sometimes alongside other forms, derived or non-derived, of the same stem. For instance, set [249] appears to be from **ʃik'-* 'to bring'.

[249]	***y 'to say, do, be'**	***ʔ 'to do'**	**Others**	
PNHEC	***ʃik' y**	***ʃik' ʔ**	**-**	'to approach'
Had	**-**	**ʃik'aaʔ-**	**-**	'approach, become close'
Kam	**ʃik' y**	**-**	**-**	'budge, move up'
Sid	**ʃik'k'i yi**	**-**	**ʃik'-, ʃik'k'iʃi**	'approach'
Ged	**ʃik'k' hiyy-, ʃikkiyya**	**ʃik'k'a**	**ʃik'-**	'approach, come close'

Where only one gloss is provided, it applies to all the stems, e.g., Sidaama and Gedeo. The Oromo parallel is *sik'a* 'move a little, drag on, push something'.

[31]	***y 'to say, do, be'**	***ʔ 'to do'**	**Others**	
PHEC	***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**	**-**	**-**	'cry'
Ged		**-**	**ood'a, ood'-**	'wail, weep, mourn', 'cry'

Sasse (1982, p. 156) also reconstructs PHEC **oo y, *oo ʔ* 'to cry' as a compound PEC **ʔoo y* meaning 'say oo' (see Oromo *oo(y)-* 'cry').

[10]	***y 'to say, do, be'**	***ʔ 'to do'**	**Others**	
PHEC	***hat't'iʃ y**	**-**	**-**	'to sneeze'
Burji	**hattiʃeys-**	**-**	**hat't'iʃee-d'-, hat'iʃ-**	'sneeze', 'cause to sneeze', 'to sneeze'
PNHEC	***hat't'iʃ y**	**-**	**-**	'to sneeze'
Had	**hat'iʃʃuy-**	**-**	**hat't'eʃ-**	'sneeze'
Kam	**hit'iʃʃu y-**	**hant'iʃʃ-aʔ, hit'iʃʃo ass-**	**-**	'sneeze'

Some languages use either support verbs to verbalize the same stem, e.g., Kambaata *hit'iʃʃu y-*, *hit'iʃʃo ass-* (Hudson, 1989, p. 138) and *hant'iʃʃ-aʔ*, 'sneeze' (Treis, p.c.).[198]

[583]	***ʔ 'to do'**	**Others**	
PHEC	***maaz ʔ, *miiz ʔ**	**-**	'to hurt'
Burji	**-**	**miid'-**	'to harm, hurt'
PNHEC	***miiz ʔ**	**-**	'to hurt'
Had	**madaʔ-**	**-**	'to wound somebody'
Kam	**maz-ees- (← *maz ʔ-is)**	**maza-ta**	'wound somebody', 'wound' (n)
Sid	**-**	**miic'i**	'hurt'
Ged	**-**	**miid'a**	'hurt, wound somebody'

[198] Treis (p.c.) does not interpret this verb as an ideophone plus a light verb, but rather as a verb whose root ends in a fossilized/frozen middle marker **aʔ-**.

Burji, Sidaama, and Gedeo parallels in other stems are influenced by the Oromo parallel *miid'a* 'hurt, be spoiled', *manc'aʔa* 'hurt'.

[673]	***y 'to say, do, be'**	**Others**	
PHEC	***ik, *y k**	**-**	'to be(come), happen'
Burji	**ih-**	**-**	'to be; become'
PNHEC	***ik-, *ikk-**	**-**	'to be(come), happen'
Had	**ikkoʔo, ih-**	**-**	'to be, become, be enough'
Kam	**ih-**	**-**	'be, suffice'
Sid	**ik-, ikk-**	**-**	'become, be, happen'
Ged	**-**	**kada**	'become, be, happen'

The geminate variant in [673] PNHEC **ikk-* 'to be(come), happen' suggests that it is a combination of the verb **y* and the agent consonant *-k*, because the meanings of PHEC **y* include 'to be'. A similar combination can be seen in the Sidaama causative verb *see-kk-* 'to adapt, make suitable', which is, according to Yri & Pepper (2019), via an intermediate **see-y-k-* and derived from the stem *see-* 'to be suitable'.

Some stems occur in compounds with only one of the two support verbs, e.g., only **ʔ* occurs with PHEC [583]. It also remains unclear if there are semantic reasons other than transitivity for the use of either verb in compounds, e.g., Kambaata *c'aʔ(a) y-* 'be(come) clean/clear (of sky)' and *c'aʔ(a) aʔ-* 'make clear/clean (by removing everything)'.

Additional examples include: PHEC [170] **sum ʔ* → PNHEC **ʃum ʔ* 'urinate', [639] PHEC **ʃaf-*, **ʃaf-ʃaf ʔ* → PNHEC **ʃaf-*, *ʃaf-ʃaf ʔ* 'to shake', PHEC [214] **t'aban ʔ, *t'aban ʔ-s-* 'to slap' → PNHEC **t'aban ʔ-s-*, and PHEC [45] **fiik' ʔ* 'to whistle'.

It should also be mentioned that almost all sources have reported the use of the two verbs as support verbs, but only after ideophones. However, it is proposed in this study that, historically, their formal use has been more than simply deriving verbs from ideophones (§3.1.2.3, §4.2.1, §4.2.3).

The transitive patterning of **ʔ* determines its distribution in denominative derivation. In §4.2.5.1, **ʔ* serves as the basis for denominal verbs with transitive or causative semantics, enabling nominal concepts to participate in transitive argument structures. The reflex **a** that appears before various extensions provides diagnostic evidence for **ʔ* as the verbalizing element (see §4.2.5.9 reanalysis framework).

The grammaticalization pathway mostly centers on **ʔ* combinations:

> Etymology of passive **-am** (§4.2.1): ***ʔ** + **-m** → ***a-m** yields the passive denominal pattern. Section 4.2.5.6 proposes that the current passive **-am** has been generalized from an older complex denominative ***ʔ+m** in order to be used with (verbal) stems not derived from nominals, establishing the grammaticalization sequence: denominal construction → general valency morphology.
>
> Alternative causative **-is** (§4.2.5.7): ***y** + **-s** → ***-is** yielding causative denominatives. The frequent attestation of causative **-is** signals a generalized structure from an older complex denominative ***y+s** used with (verbal) stems not derived from nominals, confirming the grammaticalization sequence: denominal construction → general valency morphology. Sidaama shows optional **-i-** deletion (§4.2.5.7) reflecting reanalysis of the reflex of the support verb as phonological epenthesis, demonstrating ongoing grammaticalization.
>
> Agentive layering (§§4.2.5.3–4.2.5.4): The combinations ***y-t-**, ***ʔ-t-**, and ***ʔ-k-** establish denominatives with agentive semantics. The preservation of these consonants in reflexive fusion (***ʔ-k-ʔ, *ʔ-t-ʔ** → **-ak'k', -it't'**, §4.2.5.5) demonstrates a layered structure, where the extended support verb occurs between the stem and the reflexive, creating ternary morphological units that subsequently fuse.

Understanding the vocalic reflexes of **ʔ* and **y* as verbalizers (rather than epenthetic vowels) is a prerequisite for interpreting the unified system presented in §4.2.5.9 and how the verbalized nominal stems participate in standard verbal morphology.

The support verbs ***ʔ** and ***y** retain some lexical autonomy in the compound patterns above. The next section (4.2.5) traces the grammaticalization of ***ʔ** into a non-autonomous verbalizer that attaches to nominal stems, triggering layered derivational extensions (***-t, *-k,** reflexive ***-ʔ**) documented in §§4.2.5.1–4.2.5.9.

4.2.5 Denominative Derivation

The primary denominative derivational strategy involves the addition of a support verb (§4.2.4) to nominal stems, optionally followed by verbal suffixation, to derive verbs from nouns. The term "denominative" refers to morphological markers deriving verbs from nominal bases (see §1.5.2).

The support verb in denominative constructions often undergoes phonological reduction, leaving behind vocalic or consonantal traces. This section argues that the vowel elements **i**, **a**, and **ee** preceding consonantal derivation markers in HEC are not epenthetic but rather grammatical reflexes of the verbalizers **y* → **i** and **ʔ* → **a**, with **ee** arising through coalescence (**a** + epenthetic **-i-** → **ee**). This hypothesis accounts for otherwise unexplained vowel variation in the denominative morphology and connects synchronic patterns to the diachronic grammaticalization of support verbs into derivational affixes.

This section argues that synchronic valency-changing morphology in HEC (passive **-am**, causative **-is**) originated historically as complex denominative constructions that underwent grammaticalization through functional expansion. The formal similarity between denominative and valency-changing patterns reflects a shared origin rather than coincidence or analogy.

The analysis proceeds through progressive stages of grammaticalization:

> Foundation (§§4.2.5.1–4.2.5.2): Simple denominatives establish verbalizer patterns.
>
> Layering (§§4.2.5.3–4.2.5.4): Agentive suffixes add a second derivational stage.
>
> Complex Combinations (§§4.2.5.5–4.2.5.7): Reflexive, passive, and causative extensions build elaborate denominal derivations.
>
> Theoretical Synthesis (§4.2.5.9): The complete framework integrates reanalysis, two-stage derivation, generalization, and unification.

The cross-HEC variation reflects different stages and rates of grammaticalization: Kambaata preserves distinct agentive **-k**/**-h**; Gedeo and Hadiyya show fossilized and partially active agentive **-t**; Sidaama exhibits analogy-driven reanalysis of the verbalizer **-i** as epenthetic; Burji shows contact-influenced innovations alongside inherited patterns.

Each subsection below documents one stage of this pathway with explicit cross-references showing interconnections.

4.2.5.1 The Verbalizer *ʔ

The verbalizer ***ʔ*** (← support verb **ʔ* ‘to do’) is attested across all five HEC languages, though it appears in varying degrees of morphological

transparency. In Burji, ***ʔ** persists mainly in fossilized forms of verbalized stems where its verbalizer origin is no longer synchronically productive (i.e., patterns that cannot freely extend to new nominal stems (§4.2.4.2)). In contrast, Kambaata, Hadiyya, Sidaama, and Gedeo exhibit more transparent uses in which ***ʔ** combines productively with nominal stems to derive transitive denominal verbs (see below and §§4.2.4.2, 4.2.4.3 for more examples).

	Nominal [145]		**Verbalized** [170]	
PHEC	***sum-, *sim-**	‘urine’	***sum ʔ**	‘urinate’
Burji	**sinʔa, ʃinʔaa**	‘urine’	**sinʔ-aaw-, sinʔ-ay**	‘urinate’
PNHEC	***ʃum-**	‘urine’	***ʃum ʔ**	‘urinate’
Had	**ʃuma**	‘urine’	**ʃum-eʔ-**	‘urinate’
Kam	**ʃuma-ta**	‘urine’	**ʃum-aʔ-**	‘urinate’
Sid	**ʃuma**	‘urine’	**ʃumaa-ʔr-, ʃuma-ʔr-**	‘urinate’
Ged	**ʃiʔna**	‘urine’	**ʃiʔn-**	‘urinate’

In Sidaama, **ʔr** is a variant of **d’**, and the verbalizer is reflected as **aa** on the stem. There is a generalization of the verbalized stem to the nominal in Burji and Gedeo, hence the **ʔ** in their nominal stem. The Gedeo and Burji parallels are influenced by Oromo (*finc’-eessa*; *finc’-aay-* ‘urinate’ and *finc’aan* ‘urine’).

The reflexive and the verbalizer ***ʔ** are formally identical. However, the verbalizer ***ʔ** is traced back to a transitive verbal source (**ʔ* ‘to do’; see §4.2.4), in contrast to the middle/intransitive function of reflexive ***-ʔ**. Since it is originally independent of the preceding stem, the verbalizer ***ʔ** may be realized with a vowel, but the reflexive cannot. Practically, the reflexive suffix can be attached to the verbalizer, even though it will only surface as a single **ʔ** because there is no geminate **ʔʔ**.

This subsection establishes the first stage of two-stage denominative derivation: A nominal stem combines with the verbalizer ***ʔ** to create a verbalized base that can subsequently take standard verbal morphology (i.e., Stage 2: extensions **-t**, **-k**, **-ʔ**, **-m**, **-s**, as documented in §§4.2.5.3-4.2.5.7). The fundamental structure is:

> Stage 1: Nominal + Verbalizer ***ʔ** → Verbalized Stem
> Stage 2: Verbalized Stem + Extension → Complex Verbalized Stem

Typically, simple denominatives in ***ʔ** (without additional extensions) yield transitive verbs with causative semantics, reflecting the inherent transitivity of

the source support verb (§4.2.4.2). These patterns contrast systematically with the verbalizer **y**, which yields intransitive outcomes reflecting the intransitive nature of the original verb (§4.2.4.1).

The vowel **a-** appearing after stems and extensions is not epenthetic but grammatical, i.e., it is the reflex of verbalizing ***ʔ**. This reanalysis (detailed in §4.2.5.9) explains otherwise anomalous patterns: why the vowel quality is **a** rather than expected epenthetic **-i-** (§3.4.1.1), why the vowel persists even where phonotactics would not require epenthesis, and why verbal derivation suffixes like ***-s** show a preceding vowel consistently in the denominative.

The examples below demonstrate simple denominatives with the verbalizer **ʔ** across the HEC languages. For layered combinations (***ʔ** + extension), see §4.2.5.3 (agentive ***-t**), §4.2.5.4 (agentive ***-k**), §4.2.5.5 (reflexive), §4.2.5.6 (passive ***-m**), and §4.2.5.7 (causative ***-s**). For the generalization pathway showing how ***ʔ+m** evolved into synchronic passive morphology, see §4.2.5.6.

4.2.5.2 The Verbalizer *y

Treis (2008, p. 155) reports the use of the support verb *y-* 'say' as a verbalizer **y** in Kambaata in constructions involving ideophones. Comparable uses can be inferred for other HEC languages based on the data discussed below:[199]

	Nominal		**Verbalized**	
Burji	**monnokkee**	'dull'	**monnokku iy-**	'to be dull'
Had	**c'alk'a**	'handsome'	**c'alk'ay-**	'become handsome'
Kam	**wojj-u**	'white'	**wojj y-**	'(be)come white'
Sid	**taʃʃo**	'pleasure'	**taʃʃi y-**	'to be happy'
Ged	**ʃik'o**	'near'	**ʃikkiyya**	'be near'

Parallel to §4.2.5.1, this subsection documents the verbalizer ***y** built on the intransitive support verb **y* 'to say, do, be' (§4.2.4.1). The same two-stage derivational structure applies:

> Stage 1: Nominal + Verbalizer ***y** → Verbalized Stem
> Stage 2: Verbalized Stem + Extension → Complex Verbalized Stem

[199] Data sources: Kambaata (Treis, 2008, p. 155), Burji (Roba Dame & Wedekind, 2008, p. 63; see Tesfaye Baye, 2015, p. 198), Sidaama (Indrias Xa'miso et al., 2007, p. 553), Gedeo (Wedekind, 2008, p. 94), Hadiyya (Ritter, 2007, p. 24).

However, the inherent intransitivity of ***y** determines the semantic outcomes: denominatives in ***y** yield intransitive, inchoative, or stative verbs, contrasting systematically with the transitive/causative patterns of ***ʔ** denominatives (§4.2.5.1).

The key pattern of the verbalizer ***y** is the vowel reflex **i** appearing before stems and extensions, which reflects ***y**, not epenthesis. This is demonstrated by its consistent **i** quality across environments (if epenthetic, it would show contextual variation), its persistence even where not required by phonotactics, and its systematic contrast with **a-** (← ***ʔ**) in parallel environments.

The verbalizer ***y** receives the causative ***-s** to create complex patterns (§4.2.5.7) where the sequence ***y+s** yields the denominative causative **-is**. Section 4.2.5.8 demonstrates, through comparative data, that the denominative exponents adhere to a structural template (verbalizer + extension) retained from PHEC. The verbalizer ***y** also appears in agentive layering: ***y-t-** combinations (§4.2.5.3) yield denominals with agentive semantics. The fusion of agentive **-t** with reflexive **-ʔ** creates glottalized forms (***yt-ʔ** → **-it't'**, §4.2.5.5) where the intervening **t** is preserved, providing evidence for a ternary layered structure: verbalizer + agentive + reflexive.

According to the proposed grammaticalization, denominative patterns generalize (§4.2.5.9); they may extend to word classes beyond prototypical nominals, including sound-symbolic ideophones. The expansion of denominative morphology from nominal bases to other categories exemplifies the grammaticalization process documented throughout §4.2.5.

4.2.5.3 Agentive Denominative *-t

The archaic agentive verbal extensions **-k** and **-t** historically served to specify an active thematic role for the subject in denominative derivations. While the modern semantics of these verbs have largely generalized into inchoative ('become') or stative ('be') meanings, the presence of these consonants remains a fossilized indicator of an earlier stage where the verbalized stem explicitly encoded subject-agentivity.

Synchronic evidence for verbalized stems with the (agentive) consonant ***-t** is chiefly found in Gedeo. The Gedeo reflex is **-at** (Eyob Kelemework, 2015, p. 196; Dawit Tilahun, 2019, p. 206) as shown in the following examples:

	Nominal		**Verbalized**	
[256]	**hammeenna**	‘rotten, bad’	**hamat**	‘slander’
[1030]	**huk’k’aʃʃo**	‘thin’	**huk’k’ata**	‘be slim’
[128]	**d’eebo**	‘thirst’	**d’eebota**	‘to thirst’

The cognates in the other HEC languages have corresponding variants of the denominative suffixes, e.g., [128] Gedeo *d’eeb-ota* ‘to thirst’: Kambaata *t’eeb-ek-*, *t’eeb-eʔ-* ‘be very thirsty’, Hadiyya *t’eeb-eʔ-* ‘to be thirsty’. The above examples have Oromo parallels: [256] *hammaat-*, *hamaa* (adj) ‘bad off, unwell’, [1030] *huk’k’at-* (adj) ‘thin, skinny’, and [128] *d’eebot-* (adj) ‘thirsty’. The Oromo parallels listed in the dictionaries (Gragg, 1982; Tilahun Gamta, 2004) are nominal, mostly adjectives, but they share the identical suffix **-at**. Oromo also has variants of the same roots with the middle suffix **-d’(d’)**.[200] The Gedeo and Sidaama parallels include Oromo loanwords, but it is unreasonable to rule out inheritance completely.

Sidaama stems ending with **-at**, e.g., [199], *ulat-* ‘smell’, [450], *saalfat-* ‘be shy’ (see *saal-* ‘shame’), and [698] *konkolati* ‘to roll’, are rare and some of them might be borrowed from Oromo, e.g., *ulat-* (n), *ulannaa* (n) ‘incensing or perfuming oneself with smoke of a plant that produces a sweet smell’, *saalfateeraa* ‘adolescence’, and *konkolaataa* (n) ‘car, vehicle (down hill)’, respectively. Sidaama and Gedeo stems with **-at** rarely overlap (but see Gedeo *ulat-* ‘smell good’).

In Hadiyya, a sequence **-V-t** appears to be deducible in some cases, based on available data in Ritter (2007).

Nominal		**Verbalized**	
waaka	‘greed’	**waak-at-**	‘be greedy’
gaga	‘self’	**gagi-itt-**	‘be selfish’
hit’e	‘meadow fescue grass’	**hit’-it-**	‘to graze’

There has been a discussion in the early literature about Gedeo **-at**. Hudson (1976, pp. 272, 277) considers it to be another variant of the Gedeo reflexive and attempts to justify its existence alongside **-ad’-** by suggesting “[p]erhaps it is a back-formation based on the construction of certain nouns with adjectival meaning with the copula -te…, e.g., furda-te ‘he is fat,’ reanalysed

[200] Oromo **-t-** could be interpreted as a variant of middle voice verb extensions in Borana (Marsabit) and Wata dialects reported by Stroomer (1987, pp. 153, 155) which have variants **-ad’d’-** (for the first-person singular) and **-at-** (elsewhere). However, the latter seems to derive an adjective in *huk’k’at-* (adj) ‘thin, skinny’ in Oromo, while the Gedeo parallel is an inchoative *huk’k’ata* ‘be slim’.

as a verbal construction, furd-at-e ‘he became fat,’ in which furd- is a basic stem, -at- a reflexive suffix, and -e the suffix of the 3rd sing.m. of the perfect”. Hayward (1984, pp. 89–90) rejects Hudson’s back-formation hypothesis. Given that the Oromo middle paradigm has a predominance of forms with **-t**, Hayward argues for an almost certain Oromo source for some of the examples cited by Hudson, supported by Hudson’s own note that Gedeo speakers are frequently bilingual, speaking Oromo and Gedeo. He also adds that “[f]rom the general standpoint of Highland Eastern Cushitic the forms with t are anomalous”.

Representing Stage 2 of the two-stage model, the agentive suffix ***-t** layers onto verbalized denominative stems created by the verbalizer ***ʔ** (§4.2.5.1) or ***y** (§4.2.5.2), yielding complex denominal verbs with agentive semantics.

> Nominal + Verbalizer (***ʔ/*y**) → Verbalized Stem
> Verbalized Stem + Agentive ***-t** → Complex Verbalized Verb

The provisional historical patterns would be: ***ʔ-t-** combinations yield transitive denominals with an agent/causer role, and ***y-t-** combinations yield denominals with an agent role but an intransitive frame.

The agentive ***-t** is not productively preserved in most of the HEC languages, but it appears in fossilized forms. Gedeo shows the clearest reflexes, while Hadiyya preserves traces in specific lexical items. Kambaata retained a distinct agentive ***-k** (§4.2.5.4), which potentially replaced or marginalized ***-t**. This distributional pattern suggests that ***-t** and ***-k** represent an earlier layer of PHEC morphology that has been largely superseded or lexicalized.

In addition to the evidence in Gedeo and Hadiyya, the preservation of ***-t** in reflexive combinations (§4.2.5.5) provides proof of layering. The fusion pattern ***y-t-ʔ** → **-it’t’** (agentive + reflexive → glottalized form) shows the agentive consonant intervening between the verbalizer and the reflexive. These ternary structures (verbalizer-agentive-reflexive) demonstrate that denominative derivation can involve multiple morphological layers that subsequently fuse into synchronic units. The preservation of **t** in reflexive contexts, even when lost elsewhere, exemplifies morphological fossilization.

There is provisional fossilized comparative evidence for the suffix ***-t** (§3.3.1.1) in the rest of NHEC, which complements the above **-t** in Gedeo and Hadiyya. However, a fossilized suffix **-s** is in competition with ***-t**, since they are both interpreted as complex denominatives that have lost the reflexes of the

preceding verbalizer ***y**, which would have otherwise prevented their fusion with the stem-final consonant.

[378]	**Fossilized -t**	**ø / -d'**	
PHEC	**-**	***dan-**	'to be able'
Burji	**-**	**dan-**	'be able'
PNHEC	***dan-t-**	**-**	'to be able'
Kam	**dand-**	**-**	'be able'
Sid	**dand-**	**-**	'be able'
Ged	**dand-**	**-**	'be able'

In set [378], the lack of final **d** in Burji compared to PNHEC **dan-t-* 'be able' might be due to the NHEC word having a fossilized denominative **(*y)-t** (assimilated in voicing to the final **n**), even though Oromo has a comparable form *dandaya* 'be able' (cf. Hadiyya *t'an-* 'be able').[201]

[29]	**Fossilized -t**	**b + t or s → f**	**ø / -d'**	
PHEC	**-**	**-**	***k'ab-**	'to hold, seize'
Burji	**-**	**k'af-**	**-**	'to have, possess, own'
PNHEC	**-**	**-**	***ab-**	'to hold, seize'
Ged	**abid-**	**-**	**-**	'hold, seize'

In this example, the verbalizer ***y** is preserved in Gedeo *abid-* 'hold, seize'; hence, **b** is not affected as in Burji.

As discussed in the chapter on the alternations at morpheme boundaries (§3.3.1.1), a suffixed ***(y)t** or ***(y)s** has conditioned the fricativization of root-final ***b** to **f**.[202] The regular intermediate would be geminate ***ff** from ***bs**, which is perceived as **f** because of the lack of voiceless **p** in these languages. Additionally, the single **f** reflex aligns with the lack of reconstructable geminate **ff**.

***b + *-t, *-s → *ff → f**

[201] See the following derived Hadiyya stem showing the array of derivational suffixes including **-tt**: *at'-itt-* (mid) 'to wear jewelry', *at'-it't'-am* (refl/pass) 'to beautify', *at'-is-* (tr) 'to make something good, to make something new, to repair', *at'-is-am-* (intr) 'to be correct', *at'-is-aan-co* (n) 'mender, repairer' (Ritter, 2007, pp. 12, 44)

[202] An alternative to the denominative (fossilized) **t** being the trigger for fricativization is the labial ***b** devoicing via the obstruent harmony hypothesis (fn. 36): ***b** → **f** when C_1 is a voiceless consonant.

For example, Kambaata has *huf-* ‘comb’ → *hup’p’-* (mid) ‘comb oneself’ and *t’oof-* ‘finish’, *t’oop’p’-* (mid) (Treis, 2022; Treis, p.c.).

4.2.5.4 Agentive Denominative *-k

The complex denominative suffix **-eek**/**-eeh** is attested only in Kambaata. Treis (2008) reports the form as **-eeh**, while Hudson (1989) transcribes it as **-eek**; the difference likely reflects analytical decisions rather than distinct morphemes.[203] The following set of data, mostly from Hudson (1989), presents Kambaata **-eek/-eeh**, along with the verbalizer variant **-e** + epenthetic **-ʔ-** where available in Hadiyya and Kambaata:[204]

	Nominal		**Denom -eek/h**	**Denom -ʔ**	
Kam	**t’eebo**	‘thirst’	**t’eebek-**	**t’eebeʔ-**	‘be very thirsty’
Had	**t’eebo**	‘thirst’	**-**	**t’eebeʔ-**	‘be thirsty’
Kam	**fooʃʃa**	‘breath’	**fooʃeek-**	**fooʃʃeeʔ-**	‘breathe’
Had	**fooʃʃa**	‘breath’	**-**	**fooʃʃeʔ-**	‘breathe’
Kam	**k’egu**	‘blood’	**k’egeek-, k’egeeh-**	**-**	‘bleed’
Had	**t’iiga**	‘blood’	**-**	**t’iigeʔ-**	‘bleed’
Kam	**maza-ta**	‘wound’	**mazeek-**		‘be wounded’
Had	**mada**	‘wound’	**-**	**madaʔ-**	‘to wound somebody’

Treis (2008, p. 272, fn. 23) notes in a brief footnote that “[t]he morpheme *-eeh* is a derivational morpheme which generates inchoative verbs on the basis of nouns; see, for instance, *gorr-ú* ‘hunger’ → *gorr-eeh-* ‘get hungry’”. Similarly, the adjective *haar-oo(-ta)* ‘new’ is related to the verb *haar-eeh-* ‘be(come) new’ but has no synchronic base form **haar-*.[205]

---

[203] According to Treis (p.c.), Hudson’s (1989) **-eek** is a misinterpretation because, in the perfect and perfective, **h** is strengthened and geminated to **kk**, which led Hudson to interpret this as a verb ending in **-k**. The verbal noun of these verbs clearly shows that the root ends in **h**, e.g., *mazeehu* ‘becoming wounded’. Within the context of this study, both analyses are acceptable because non-initial ***-k** becomes **-h** in all HEC, including Kambaata (§2.1.5.2 (PHEC *h)), while non-initial ***-h** has been lost throughout HEC (§3.2.2). Therefore, it is an analytical decision to represent the more abstract **-eek** or the more concrete **-eeh** – a useful disagreement for historical analysis.

[204] Data sources: Kambaata (Hudson, 1989, p. 82; Treis, 2008; Alemu Banta, 2016), Hadiyya (Hudson, 1989; Ritter, 2007).

[205] Treis (2008, p. 170) says, referring to examples such as the ones involving the inchoative, that “[t]he derivational formative by which a noun is created on the base of a certain verb is not predictable but conventionalized. So far hardly any conditioning factors for the choice of one or the other derivational pattern could be determined[…]. Nothing helps us to predict that nouns corresponding to the verbs [e.g., the de-nominal verbs *maz-eeh-* ‘be wounded’ ← *maza-ta* ‘wound’ and *gorr-eeh-* ‘be(come) hungry’ ← *gorr-u* ‘hunger’] share a stem with them (i.e. are “zero-derived”), are derived from them with a certain formative, or are themselves the base for the verb”. The answer to that

Further examples include:

> Kam *barr-eek-* ‘be late (of day)’, Burji *orr-aaw-* ‘be late (of day)’ (intr), *orr-aa-s-ad’-* ‘be late in morning (of person)’,
>
> Kam *gorru* ‘hunger’, *gorr-eeh-*, *gorr-ees-* ‘be hungry’

Hadiyya and Kambaata **e?** reflect the corresponding denominative, with a raised verbalizer vowel **-a** and an inserted glottal stop. The glottal stop is epenthetic after lenition (***-ek** → ***eh**) and the elision of an intermediate ***h** (see §3.2.2). Hadiyya shows a similar pattern of epenthesis, replacing an underlying consonant (see §3.3.4.1). In other words, an intermediate reflex **ah** of the verbalizer ***?** and the lenited agentive ***-k** yielded a compensatorily doubled (and optionally raised) vowel **ee**, resulting in **ee** and a recently suffixed **-k** – e.g., Hadiyya *mada?-* ‘to wound somebody’ (tr), verbalized with ***?**, not the agentive **-e-?-**, and Kambaata *mazeek-* ‘be wounded’ with agentive layer **-eek**. In this view, the variants with **e?** (underlying **eh**) in the above examples in Kambaata and Hadiyya reflect an intermediate stage before the addition of **-k**, e.g., Kambaata *t’eebe?-* ‘be very thirsty’, Hadiyya *t’eebe?-* ‘be thirsty’, and with support verb + **-k** in Kambaata *t’eebek-* ‘be very thirsty’.

An alternative hypothesis traces **-eek** to the coalescence of the verbalizer **a** with the verb [673] PHEC **ik-* → PNHEC **ik-*, **ikk-* ‘to be(come), happen’ (**a-** + **-ik** → **-eek**). However, no independent HEC evidence for **ik-* as a verbalizer component has been identified, and the simple morphophonological account (***-k** lenition + Kambaata compensatory doubling / Hadiyya epenthesis) requires fewer assumptions. The **ik-* hypothesis remains tentative pending further investigation.

In addition to the agentive suffix ***-t** (§4.2.5.3), the agentive ***-k** (realized as **-eek/-eeh/-e?** reflecting regular sound changes) represents Stage 2 of the two-stage model layering onto verbalized denominative stems created by the verbalizer ***?** (§4.2.5.1). It serves the same derivational function as ***-t** (§4.2.5.3) but remains productive in Kambaata. Its structure is parallel to ***-t** as follows:

---

puzzle, at least for the inchoative, would be that the nouns themselves are the base for the verb (see §4.2.3).

Nominal + Verbalizer (***ʔ**) → Verbalized Stem
Verbalized Stem + Agentive ***-k** (or [673] PHEC **ik-*) → Complex Verbalized Verb

Similar to Gedeo **-at**, **-eek/-eeh** appears robustly in Kambaata with inferable parallels in Hayiyya. This distribution pattern indicates functional replacement where **-k** occupies the derivational space that **-t** held in earlier stages. Alternatively, languages may have generalized one of the two, reducing the duality of the archaic morphology (assuming an earlier agent gender specification of the two).

Like ***-t**, the agentive ***-k** fuses with reflexive **-ʔ** to create a glottalized reflexive suffix via ***ʔ-k-ʔ** → **-ak'k'** in Kambaata and possibly Hadiyya (§4.2.3.2). This parallel fusion demonstrates that ***-k** follows the same structural template as ***-t**, occupying the middle position in ternary verbalizer-agentive-reflexive structures (§4.2.5.5). The replication of fusion patterns across different agentive suffixes suggests that the layered derivational template and the morphemes are inherited.

The distribution of agentive ***-t** (productive in Gedeo and marginal in Hadiyya) versus ***-k** (productive in Kambaata and inferable in Hadiyya) suggests functional replacement rather than phonological alternation: Both suffixes occupy the same derivational slot (post-verbalizer, pre-reflexive) and encode similar agentive/inchoative semantics, but show complementary language distribution. The shared Hadiyya and Kambaata fusion pattern (***-t-ʔ** / ***-k-ʔ** → glottalization, §§4.2.5.3–4.2.5.5)[206] suggests both suffixes are integrated into the same inherited template structure, favoring the functional replacement scenario over independent innovation. Subgrouping implications are examined in §5.3. Therefore, the agentive suffixes ***-t** and ***-k** are reconstructable to PNHEC, since Burji shows no reflexes.

The coexistence of scattered reflexes of ***-t** and ***-k** indicates that morphological replacement and erosion can occur at the level of individual derivational slots while preserving overall structural templates, productivity can shift between competing exponents of the same derivational function, and synchronic variation in NHEC reflects diachronic layering of ***-t** and ***-k**. This pattern of template preservation extends to the generalization process in

[206] Plazikowsky-Brauner (1957, p. 29) records a non-glottalized verbal reflexive suffix **-ak** for Hadiyya (§4.2.3.2).

§4.2.5.9, i.e., denominative templates (verbalizer + extension) generalize to valency-changing functions, while individual morphemes may be reanalyzed.

4.2.5.5 The Reflexive Denominatives

Several HEC grammars (Eyob Kelemework, 2015, p. 196; Treis, 2008, p. 271; Kawachi, 2007, p. 134; Tesfaye Baye, 2015, p. 197) report formal overlap between reflexive markers (§4.2.3) and the reflexive denominatives. The same morpheme (e.g., **ʔ**) appears on verbal stems and on nominal stems. This highlights the fuzzy formal and functional boundaries between markers of reflexives and denominatives in HEC. This section treats the nominal-stem cases as denominative, acknowledging their formal identity with the reflexives (§4.2.3), and proposing their historical trajectories.

	Form	**Nominal**		**Verbalized**	
Kam	**-ʔ**	**gaan-a**	‘fat’	**gaaʔn-**	‘be(come) fat’
		nubaacc-u	‘old’	**nubaap’p’-**	‘be(come) old’
	-it’t’	**gabbanc-u**	‘short’	**gabb-oo-t’t’-**	‘be(come) short’
Had	**-it’t’**	**hoffi**	‘small’	**hoff-ii-t’t’-**	‘become poor’
		gunda	‘short’	**gund-i-t’t’-**	‘shorten’
Ged	**-ʔ**	**haruma**	‘short’	**haruʔm-**	‘become short’
	-ed’	**k’eerra**	‘tall’	**k’eerr-ed’-**	‘become tall’
Sid	**-ʔ**	**duumo**	‘red’	**duuʔm-**	‘become red’
	-i-d’	**waajjo**	‘white’	**waajj-i-d’**	‘become white’
Burji	**-ʔ**	**gaani**	‘big’	**gaanʔ-**	‘become big’
	-ad’	**c’arbaá**	‘wet’	**c’arbaad’-**	‘become wet’
	-eed’	**duwwaaga**	‘red’	**duwweed’-**	‘become red’

The above examples show a strong resemblance between reflexive and denominative markers. However, they are analyzed as denominatives, and where they are identical to the reflexives, their denominative form is the older formation, which has been generalized to the reflexive.

Gedeo and Sidaama use several suffixes to derive intransitive denominal verbs from nominal stems. In Gedeo, these include **-ʔ**, **-at**, and **-ed’**; in Sidaama, **-ʔ** and **-i-d’**.[207] According to Eyob Kelemework (2015, p. 197), the

[207] Other forms or allomorphs are listed in the sources (e.g., Sidaama: **-ʔr**, **-p’**, **-t’** (Kawachi, 2007, p. 134)), but they are only products of phonological or morphophonemic processes.

distribution of these markers appears to be lexically determined rather than following a predictable phonological or semantic rule in Gedeo.[208]

Tesfaye Baye (2015, p. 183) specifies that **-eed'** with a double vowel is only used as a denominative in Burji. That vowel doubling is due to the coalescence of the two segments: a reflex of the verbalizer ***y** (§4.2.4.1) and the initial vowel of the reflexive suffix **-ad'** (§4.2.3.4). Additionally, Tesfaye Baye does not list the glottal stop as a denominative. However, there are a small number of examples with glottal stop **ʔ**, identified as middle forms in Roba Dame & Wedekind's (2008) Burji Dictionary. These are, however, better analyzed as verbalizers (even if they are underlyingly extended ***ʔ-ʔ**), such as *gaani* 'big' → *gaanʔ-* 'to become big, grow' with the causative **-s** replacing the glottal stop in *gaans-* (caus) 'to make big, enlarge'; similarly in *ʃaalla* 'shame' → *ʃaalʔ-* 'to be ashamed' compared with *ʃaallis-* (caus) 'to ashame'. The denominative is also formed by utilizing the verbalizer **y** in Burji *d'iid'aa* 'jealousy' → *d'iid'iy-* 'to become jealous' and before the causative (→ **i**) in *d'iid'is-* (caus) 'to make jealous' (Roba Dame & Wedekind, 2008).

Kambaata **-it't'** and **-ak'k'** are attested but not frequent as denominatives, and there is possibly even a rare form **-t** (§4.2.5.3), e.g., *gabban-c-u* 'short' → *gabb-oo-t't'-* (intr) 'be(come) short', *gabboot-is-* 'shorten' with **-i-** coalescence with a vowel-based reflex of the verbalizer ***ʔ** (§4.2.4.2). Due to their scarcity, they are not mentioned as independent suffixes in any of the previous literature. Likewise, Treis briefly mentioned the denominative **-eeh** (transcribed as **-eek** in Hudson (1989)) and **-ʔ** only as a middle marker obligatorily occurring on some verbs (Treis, 2008, p. 272).[209]

Hadiyya has further denominative suffixes discussed in the previous subsections (§§4.2.5.1–4.2.5.4); however, they are not documented in the literature and are mainly inferred from the available data.

The combination of verbalized stems with reflexive morphology reveals evidence for a layered morphological structure coming about by means of fusion processes. The standard reflexive patterns (baseline from §4.2.3) is the simple attachment of the basic reflexive ***-ʔ** (§4.2.3.1) to verbal stems with the

[208] Some Gedeo examples exhibit double derivation with the non-glottalized **-t** and the glottalized **-t't'** added to the adjective form: *hortaat't'e* (adj) 'fertile (of woman)', *hor-em-* (intr) 'breed, multiply', *hor-siis* 'breed' (tr) and *ildaat't'e* (adj) 'fertile (of woman)', *il-* (tr) 'give birth, beget'.

[209] According to Treis (2008, p. 272), crucially, several inchoative verbs derived from adjectives are "obligatorily middle-marked" in Kambaata.

expected regular alternation process in NHEC (§§3.3.1, 3.3.2). The complex denominative reflexives, on the other hand, include the support verb - agentive - reflexive ternary structures, and again, exhibit the NHEC alternation process affecting the intervening consonants.

The defining characteristic of denominative reflexives in Kambaata and Hadiyya is the preservation of agentive consonants (***-t/*-k**) in reflexive structures even when these consonants are lost or marginal elsewhere. The first pattern (***y-t-ʔ** → **-it't'**) is comprised of the verbalizer *y (§4.2.5.2) + agentive ***-t** (§4.2.5.3) + reflexive ***-ʔ** (§4.2.3.3, mostly Hadiyya). The agentive **-t** is fossilized in fusion, in addition to its preservation in non-reflexive contexts, in Gedeo and Hadiyya. The second pattern (***ʔ-k-ʔ** → **-ak'k'**) includes the verbalizer ***ʔ** (§4.2.5.1) + agentive ***-k** (§4.2.5.4) + reflexive ***-ʔ** (§4.2.3.3, mostly Kambaata). The agentive **-k** preserved in Kambaata and traceable in Hadiyya (§4.2.5.4) shows the same structural template as ***-t** in the reflexive.

The preservation of agentive consonants in fusion with the reflexive proves a ternary morphological structure. Denominative derivation involves three layers (support verb + agentive + reflexive), not just two (stem + reflexive). The agentive suffix slot is between the verbalizer and the reflexive. Morphological fossilization is evident in the preserved agentive ***-t** (§4.2.5.3), which is marginally inferred from the reflexive fusion. The same NHEC fusion pattern (verbalizer-agentive-reflexive → glottalized agentive consonant) applies to both ***-t** and ***-k**, proving that the derivational template is maintained regularly.

Non-denominative verbs take simple reflexive ***-ʔ** without intervening consonants: verbal stem + ***-ʔ** → reflexive verb (no suffix-internal complexity). The presence vs. absence of intervening agentive consonants thus serves as a diagnostic for the determination of denominative origin: Glottalized reflexives with **-ak'k'/-it't'** patterns betray underlying denominative structure even when synchronic speakers no longer analyze the forms as compositional (§4.2.5.9).

Over time, fused forms like **-it't'** undergo semantic specialization: The relationship between derived and base forms becomes less transparent. For example, a verb like Hadiyya *sigg-it't'-* 'to cool down' may develop idiosyncratic meanings beyond predictable "agent-reflexive" semantics. This semantic drift, coupled with formal fusion, obscures the underlying morphological complexity for synchronic speakers while preserving structural evidence for diachronic analysis.

The fusion processes documented here exemplify broader patterns in the grammaticalization pathway: Morphologically complex forms (ternary structures) undergo phonological fusion, fused forms develop specialized semantics diverging from compositional origins, original morphological boundaries become opaque to synchronic analysis, and structural remnants (preserved consonants) persist as diachronic traces. These same processes operate in the generalization of complex denominatives (***ʔ+m**, ***y+s**) into synchronic valency morphology (**-am** passive, **-is** causative): formal complexity → fusion → semantic specialization → categorial expansion.

The following is a summary of the complex denominatives/reflexives and the historical sources of their exponents discussed in this section:

Verbalizer	Agentive ***-t**	Agentive ***-k**	Reflexive **-(V)d'**
***ʔ**	**-at** (§4.2.5.3)	**-eek/-eeh** (§4.2.5.4)	**-eed'** (§4.2.3.4)
		↓	
	-	**-ak'k'** (§4.2.3.2)	**-ad'/-ed'** (§4.2.3.4)
***y**	**-it** (§4.2.5.3)	-	**-id'** (§4.2.3.4)
	↓		
	-i-t't' (§4.2.3.3)	-	-

The innovated reflexive **-(V)d'** is added as it is after the verbalizer in Burji, Gedeo, and Sidaama. The reflex of the verbalizer is more transparent in the double vowels of Burji **-eed'** and maybe Sidaama **-id'**. It should be noted that some of the instances that might seem to be simple verbalizers – i.e., they show the verbalizer **ʔ** only – might underlyingly be comprised of two glottal stops: the verbalizer **ʔ** plus the reflexive **-ʔ**.

4.2.5.6 The Passive Denominative

Gedeo and Burji show **-V(m)m** denominative forms extended with the passive ***-am**: Gedeo **-om** (Eyob Kelemework, 2015, p. 197) and Burji **-oom**, **-umm** (Tesfaye Baye, 2015, p. 198).[210]

[210] Data sources: Gedeo (Wedekind, 2008), Sidaama (Kawachi, 2007), and Burji (Tesfaye Baye, 2015; Roba Dame & Wedekind, 2008).

	Nominal		Verbalized	
Ged	**araarsaa**	‘peacemaker’	**araarema**	‘be reconciled’
	hiyyo	‘poverty’	**hiyyom**	‘become poor’
Sid	**araara**	‘reconciliation’	**araaram-**	‘become reconciled’
	k’arra	‘trouble’	**k’arram-**	‘be in trouble’
Burji	**haraaroo**	‘reconciliation’	**haraarmad’-**	‘become reconciled’
	hiyyuntaa	‘poverty’	**hiyyumm-**	‘become poor’
Had	**araassaanco**	‘mediator’	**araaram-**	‘make peace’ (intr)
	k’uunk’a	‘egg’	**k’unk’am**	‘be round’

There is a parallel **-am** in Sidaama even though Kawachi (2007, p. 335) does not recognize it as a denominative, instead noting that “[t]here are a few verbs that always contain the passive suffix and cannot be used as verbs without it, though they usually have related noun forms: k’arr-am- ‘to be in trouble’ (k’arra ‘problem, trouble’), araar-am- ‘to become reconciled’ (araara ‘reconciliation’, araar-s- ‘to reconcile’)”.

In Hadiyya, Tadesse Sibamo (2015, p. 145) only reports the use of **-aam** for deriving adjectives from nouns. However, the dictionary entries in Ritter (2007) include numerous denominal forms (marked as passive/reflexive) with **-am** matching the denominative in the other languages. Therefore, this study posits the denominative function for Hadiyya, as well.

The combination of denominative stems with the passive morpheme ***-m** establishes the etymological foundation for synchronic passive **-am**. The analysis here is crucial for understanding the generalization pathway detailed in §4.2.5.9.

The synchronic passive suffix **-am** across the HEC languages derives historically from a complex denominative construction: ***ʔ** + middle/passive ***-m** → ***a-m** → **-am**. At the first stage, the verbalizer ***ʔ** (§§4.2.5.1, 4.2.4.2) converts the nominal stem into a verbal base. At the second stage, middle/passive ***-m** (see §4.2.1 for basic passive patterns) attaches to the verbalized stem, yielding a denominalized passive with vowel **a** as a reflex of the verbalizer ***ʔ**, not as an epenthetic insertion. The complex denominative ***ʔ+m** originally derived passive verbs from nominal bases, enabling nominal concepts to participate in passive argument structures. The examples above illustrate this denominal pattern. The functions are compositional at this stage: denominal + passivization = passive denominative.

The comparative data in §4.2.1 suggest that the current passive **-am** is generalized from an older complex denominative ***ʔ+m** used with stems that

have not been derived from nominals. This establishes the grammaticalization sequence in two stages: original (a) nominal + ***ʔ+m** → denominal passive (restricted to verbalized stems), then generalized (b) any verbal stem + **-am** → passive (expanded to all transitive verbs). The generalization represents functional expansion, i.e., a derivational pattern originally restricted to denominative contexts extends to apply productively to any eligible verbal stem. The formal structure (**-am**) is preserved while the functional domain expands categorially.

Examples in §4.2.5.8 and §4.2.1 present comparative evidence for the diverse selections of the denominatives used on the same nominal base across HEC.

Understanding **a** as the reflex of the verbalizer ***ʔ** (rather than epenthetic) is essential for this interpretation. The reanalysis framework in §4.2.5.9 shows that what have traditionally been treated as epenthetic vowels in denominative constructions are in fact reflexes of these support verbs. Explaining the appearance of non-epenthetic vowels like **a** without recognizing ***ʔ** → **a** would obscure the proposed ***ʔ+m** → **-am**.

Burji and Gedeo show an alternative pattern: ***ʔ + w** (see the rounded vowels) → inchoative marking. This is attributed to Oromo contact (as a source for the **-w** based inchoative), thus representing local innovation via contact (Oromo influence on Burji/Gedeo) and a parallel structure: ***ʔ** + extension following the inherited template with subsequent concatenation of **-m**. Irregular rounding in the denominative construction due to the labial **-m** may alternatively explain the rounding.

The etymology of the passive morpheme **-am** exemplifies the central claim of §4.2.5, i.e., synchronic valency morphology originates in denominative constructions that generalize. The same patterns recur in causative **-is** (§4.2.5.7) and are synthesized in the theoretical framework presented in §§4.2.5.8–4.2.5.9 .

4.2.5.7 The Causative Denominative

A denominative derivational suffix -**is** is formally similar to the causative **-is**. Unlike the causative, where **-i-** is epenthetic, the denominative includes a reflex of the verbalizer ***y**, which surfaces as **i** or is deleted between consonants.

	Nominal		**Verbalized**	
Burji	**haraaroo**	‘reconciliation’	**haraar-s-**	‘mediate, reconcile’
Had	**araaro**	‘peace’	**araar-, araass-**	‘to reconcile’
Kam	**araaruta**	‘intervention’	**araar-s-**	‘mediate, intervene’
Sid	**araara**	‘reconciliation’	**araar-(i)-s-**	‘mediate, reconcile’
Ged	**araarsaa**	‘peacemaker’	**araar-s-**	‘mediate, reconcile’

Sidaama’s optional **i** in the verbalized *araar-(i)-s-* ‘mediate, reconcile’ reflects the verbalizer **y**, which has otherwise been lost throughout the HEC verbalized stems. Its elision in this environment may be motivated by analogy with the causative allomorph **-s** after single sonorants (cf. 3.1.2.2): Once **i** was reinterpreted as epenthetic, it could be deleted, since **-i-** is not obligatory between a stem-final single sonorant and **-s**.

The other verbalizer **ʔ** has also contributed to another common denominative **-ees(s)**, arguably through coalescence via ***ʔ-is** → ***a-is**. In this case, the causative variant **-is** (with epenthetic **-i-**) is preceded by the verbalizer **a-**.[211]

	Nominal		**Verbalized**	
Burji	**bayyinna**	‘distance’	**bayyeesa**	‘keep away’
Had	**eddecca**	‘clothes, clothing’	**edd-ees-**	‘to clothe somebody’
Kam	**kaakk**	‘cackling sound’	**kaakk-ees-**	‘cackle (of a hen)’
Sid	**wodan-**	‘heart’	**wodan-eess-**	‘make someone thoughtful’
Ged	**mada**	‘wound’	**madeess-**	‘wound’ (caus)

The geminated consonant **ss** in Sidaama and Gedeo is puzzling and might require a different analysis than the one posited here if they are the result of fusion of the causative **-s** with another consonant or extension with an additional causative (**-ees+-s**).

The following sets include two verbalized and extended forms of a lost source **ħuy-* ‘full’, using the complex (extended) forms of the denominative **ʔ-m** and **ʔ-s**:

[211] Data sources: Hadiyya (Ritter, 2007), Kambaata (Alemu Banta, 2016), Sidaama (Yri & Pepper, 2019), Gedeo (Eyob Kelemework, 2015).

	Denom ʔ-m- [520]		**Denom ʔ-s-** [607]	
PHEC	***ħuy** → ***ħuy ʔ-m-**	‘be(come) full’	***ħuy ʔ-s-**	‘to fill’
Burji	**huuʔ-, huum-**	‘be(come) full’	**huuss-, huus-**	‘fill’
PNHEC	***ħuy ʔ-m-**	‘be(come) full’	***ħuym-s-**	‘to fill’
Had	**woʔm-**	‘be(come) full’	**wanʃ-, wonʃ-**	‘fill’
Kam	**wiim-, yiim-**	‘be(come) full’	**wiinʃ-, yiinʃ-**	‘fill’
Sid	**woʔm-**	‘be(come) full’	**wonʃ-, woʔm-is-**	‘fill’
Ged	**woʔm-**	‘be(come) full’	**wonʃ-, wanʃ-**	‘be full, fill’

PHEC **ħuy-* ‘full’ becomes **ħuy ʔ-m-* ‘be(come) full’ and **ħuy ʔ-s-* ‘to fill’, while PNHEC **ħuym-s* ‘to fill’ uses the verbalized stem (instead of the root **ħuy-*) as a derivational base, upon which it adds the causative **-s** without a need for a verbalizer. Sidaama even generalized the verbalized *woʔm-* ‘be(come) full’, so it uses the causative with epenthesis (**-i-s**) in *woʔm-is-* ‘fill’ because of the stem-final cluster (see §4.2.2.2). Kambaata *wiim-*, *yiim-* ‘be(come) full’ and *fiil-* ‘unravel’ might have lost the fossilized glottal stop and compensated for that loss by vowel doubling. In addition to the above, see Burji *huuwʔa* ‘to be filled up, be full of’ (Getahun Addo, 2017, p. 478) and Gedeo *uwwa* ‘full’ (Dawit Tilahun, 2019, pp. 59, 55), with the assimilation of stem-final **y** to **w** after **u**.

The combination of denominative stems with the causative **-s** reveals multiple etymological pathways that converge on formally similar surface patterns. The analysis establishes critical foundations for the generalization framework in §§4.2.5.6, 4.2.5.7, and 4.2.5.9.

Two distinct patterns emerge in the causative denominative. First, **-is** with ***y** reflex (← ***y+s**) via nominal + ***y** (intransitive support verb) → verbalized stem, verbalized stem + causative ***-s** → denominative causative. This is the transparent denominative causative with the verbalizer preserved as **i**, distinguishing it from the simple causative **-is** on non-derived verbs (where **-i-** is epenthetic, see §4.2.2.2). Second, alternative causative **-ees** via coalescence (← ***ʔ+is** → ***a-is** → **-ees**) via nominal + ***ʔ** (← transitive support verb) → verbalized stem + causative **-is** → denominative causative. This pattern shows the verbalizer ***ʔ** (§4.2.5.1) combining with the causative suffix **-is**.

The denominative causative patterns documented here provide the etymological source for the synchronic causative morphology. Sections §4.2.2.2 and §4.2.5.7 suggest that the denominative/derived support verb **-is**

from ***y+s** has been generalized to the standard causative derivation. This means that the original function of the denominative causative (restricted to verbs derived from nouns) has been generalized to the general causative (which applies to any verbal stem). The formal **-is** structure is maintained through functional expansion, and the original verbalizer reflex (**i** ← ***y**) is sometimes dropped, enabling analogical processes like Sidaama deletion or reinterpretation as causative **-s** (§4.2.2.1).

The generalization pathway for causative **-is** parallels that of the passive **-am**: in the passive, ***ʔ+m** (denominative) → **-am** (general valency marker) and in the causative, ***y+s** (denominative) → **-is** (general valency marker). Both represent functional expansion from denominal derivation to general valency alternation, preserving formal structure (***ʔ** → **a,** ***y** → **i**) while expanding categorial distribution.

Synchronically, the two sources of surface **-is** in the HEC languages are the denominative **-is**, which contains verbalizer reflex ***y** → **i** (grammatical), and the causative **-i-s** with epenthetic **-i-** (phonological). The denominative cannot delete **-i-** without destroying derivational transparency; however, in the causative, **-i-** is subject to the epenthesis rules set in §3.1.2.2.

The deletion of the deverbal **y/i** in denominative causatives represents the collapse of this distinction. As **-is** generalizes, the denominative pattern is reanalyzed as phonological, enabling both patterns to converge formally and functionally.

The following summary presents the historical sources of the complex passive and causative denominatives:

Verbalizer	Passive ***-m**	Causative ***-s**
ʔ	**-am** (§4.2.1)	**-ees** (§4.2.2.2)
y	-	**-is** (§4.2.2.1)

The overall generalization pathway is synthesized in §4.2.5.9.

4.2.5.8 Comparative Data

This subsection presents comparative data demonstrating the complex denominative constructions discussed in the previous section.

The following examples illustrate the role of different denominative morphemes in deriving verbs from nominals, as documented in the

comparative database. The glosses are provided directly beneath each lexeme. The verbalizers in the first row of the table have different vocalic reflexes when followed by one of the extension consonants (**-t**, **-k**, **-s**, **-m**). Otherwise, they may surface as the consonants **ʔ** or **y** when not followed by any of those consonants, as illustrated in [392] and [583].

Example [426] below presents comparative data for the reconstruction of denominative derivations. Using PHEC [426] **ħaraar y-s-, *araar y-s-* as a diagnostic case, the table tracks the selection of verbalizing strategies across the languages.[212]

[426]		**Verbalizer ʔ**	**Verbalizer y**	**Nominal Stem**
Burji	**-s**	-	*haraar-s-* 'mediate, reconcile'	*haraaroo* 'reconciliation'
Had	**-s**	-	*araass-* 'to reconcile'	*araaro* 'peace'
Kam	**-s**	-	*araar-s-* 'mediate, intervene'	*araaru-ta* 'intervention'
Sid	**-s**	-	*araar-i-s-*, *araar-s-* 'to reconcile'	*araara* 'reconciliation'
Ged	**-s**	-	*araar-s-* 'mediate, reconcile'	*araarsaa* 'peacemaker'
Had	**-m**	*araar-am-* 'make peace' (intr)	-	-
Sid	**-m**	*araar-am-* 'to become reconciled'	-	-
Ged	**-m**	*araar-em-* 'be reconciled'	-	-

The data distinguishes between the causative denominative and passive denominative, e.g., in Hadiyya. While ***ʔ** yields regular **a**/**e**, the vowel reflexes of the verbalizer ***y** vary between the optional remnant **i** in Sidaama and loss in the rest of HEC, as discussed in §4.2.5.7.

[212] Data sources: All languages (Hudson, 1989), Sidaama (Kawachi, 2007), Hadiyya (Ritter, 2007), Gedeo (Wedekind, 2008; Gasparini, 1995, p. 95), Kambaata (Treis, 2008; Alemu Banta, 2016).

[392]		Verbalizer ʔ	Verbalizer y	Nominal Stem
Had	**ø**	-	**lik'aayy-** 'to borrow something'	**lik'aayyo** 'loan, debt, lending'
Burji	**-t**	**lik'eesad'** 'borrow'	**-**	**lik'ee** 'loan'
Had	**-t**	**-**	**lik'aayyit't'-** 'to borrow'	**-**
Sid	**-t/d'**	**-**	**lik'-iʔr-is-** 'to lend (sth. cannot be returned as it is)'	**-**
Ged	**-t**	**lik'eesseda** 'to borrow'	**-**	**-**
Burji	**-s**	**lik'ees-** 'lend (money)'	**-** -	**-**
Had	**-s**	**-**	**lik'aayyis-** 'lend'	**-**
Kam	**-s**	**leʔees-** 'lend of what cannot be returned'	**-**	**-**
Sid	**-s**	**-**	**lik'iiss-** 'lend'	**-**
Ged	**-s**	**lik'eess-** 'lend (money)'	**-**	**lik'a** 'debt, borrowed property'

The stem **lik'-* 'to lend' shows multiple derivational strategies across languages, including simple denominative in Hadiyya *lik'aayy-*, reflexive denominative in Burji *lik'eesad'-* and Hadiyya *lik'aayyit't'-*, and causative denominative in Sidaama *lik'iiss-* 'lend'.

[285]		Verbalizer ʔ	Verbalizer y	Nominal Stem
Burji	**-s**	**-**	**balʔ-is-** 'widen'	**balʔaneega** 'wide' (adj.m)
Ged	**-s**	**-**	**bad'd'-is-** 'widen'	**baʔla** 'wide, broad'
Burji	**-t/d'**	**balʔad'-** 'to become wide, broad'	**-**	**-**

[128]		Verbalizer ʔ	Nominal Stem
Had	**ø**	**t'eebeʔ-** 'to be thirsty'	**t'eebo** 'thirst'
Kam	**ø**	**t'eebeʔ-** 'be very thirsty'	**-**
Burji	**-t/d'**	**d'aybood'-** 'to become thirsty'	**d'ayboo** 'thirst'
Ged	**-t**	**d'eebota** 'thirst'	**deebo** 'thirst'
Kam	**-k**	**t'eebek-** 'be very thirsty'	**-**

[583]		Verbalizer ʔ	Verbalizer y	Nominal Stem
Had	**ø**	**madaʔ-** 'to wound somebody'	**-**	**mada** 'wound'
Sid	**ø**	**miic'i** 'hurt'	**-**	**-**
Ged	**ø**	**miid'a, madoʔ-** 'hurt, wound somebody', 'wound'	**-**	**mada, madoʔa** 'wound'
Kam	**-k**	**maz-eek-, maz-eeh-** 'be wounded'	**-**	**maza-ta** 'wound'
Kam	**-s**	**maz-ees-** 'wound somebody'	**-**	**-**
Sid		**-**	**mad-iiss-** 'to wound'	**-**
Ged	**-s**	**madeess-** 'wound'	**-**	**-**

The comparative data above demonstrate that denominative derivation in the HEC languages involves a range of suffixes, including causative, passive, and reflexive markers. The distribution of these suffixes and the reflexes of the verbalizers highlight the complex historical layering. The evidence from Burji, Gedeo, Sidaama, and Kambaata supports the view that denominal derivation is both morphologically diverse and sensitive to language-specific changes. Further research is needed to refine the criteria for distinguishing the function of each denominative in HEC verb formation.

4.2.5.9 Summary and Remarks

This section synthesizes the denominative analysis presented in §§4.2.5.1–4.2.5.8, articulating the complete theoretical framework that accounts for the

formal, arguably functional, and distributional properties of HEC denominative and valency-changing morphology. They are summarized in five key insights.

Reanalysis Framework

The analysis establishes that vowels appearing before verbal extensions are not phonologically epenthetic but grammatically meaningful, serving as reflexes of verbalizers: **i** ← ***y** (from the intransitive support verb, §§4.2.4.1, 4.2.5.2), **a** ← ***ʔ** (from transitive support verb, §§4.2.4.2, 4.2.5.1), and **ee** ← coalescence of ***ʔ** + **i** → **-ees** (§4.2.5.7) or extended ***y**.

The denominative system displays a clear structural pattern in which support verbs function as core elements that enable category change. This pattern supports the analysis that what have traditionally been treated as epenthetic vowels in denominative constructions are in fact reflexes of these support verbs, explaining the appearance of non-epenthetic vowels like **a** and **ee**.

The epenthesis analyses predict that vowels are inserted to satisfy phonotactic constraints (§§3.4.1, 3.1.2.2, 3.1.1.5). They should show contextually predictable quality (e.g., high vowel **-i-** in most HEC languages). They should appear only where phonotactics require them (e.g., for breaking illegal consonant clusters). They should exhibit regular reflexes in quality. However, denominative vowels show non-predictable quality, with **a** appearing where epenthesis would predict **-i-**; non-required contexts, with vowels persisting even where phonotactics allow consonant sequences; and systematic distribution, with **i** correlating to ***y** for cases of intransitive semantics and **a** correlating to ***ʔ** for cases of transitive semantics. This distribution is explainable if the vowels are grammatical (verbalizer reflexes) rather than phonological (epenthesis).

Treating **i/a** as grammatical enables the reconstruction of denominative etymology even in synchronically opaque forms. For example, passive **-am** transparently derives from ***ʔ+m** once **a** is recognized as grammatical (from ***ʔ**), whereas its analysis as epenthetic would obscure this etymology.

Two-Stage Derivation Model

Denominative formation in PHEC involved a two-stage process:

> Stage 1 – Verbalization: Nominal Stem + Verbalizer (***y** or ***ʔ**) → Verbalized Stem

The verbalizer enables category change from noun to verb, and the transitivity of the source support verb determines the argument structure of the derived verb, i.e., broadly ***y** with intransitive/inchoative meanings and ***ʔ** with transitive/causative meanings.

> Stage 2 – Extension: Verbalized Stem + Verbal Morphemes **(*-t, *-k, *-ʔ, *-m, *-s**) → Complex Deverbalized Stem

The standard verbal extensions that attach to the verbalized stems are agentive (§§4.2.5.3–4.2.5.4), reflexive (§4.2.5.5), passive (§4.2.5.6), and causative (§4.2.5.7). The result is ternary structures like ***y-t-ʔ** (verbalizer + agentive + reflexive) → synchronic **-it't'** via fusion.

The preservation of intervening consonants in reflexive fusion (§4.2.5.5) proves layering: forms like **-it't'** (§4.2.3.3) and **-ak'k'** (§4.2.3.2) show agentive ***-t** and ***-k** between a support verb (***y** or ***ʔ**) and the reflexive morpheme (***-ʔ**), demonstrating a three-layer structure rather than a binary stem + reflexive.

The two-stage model explains why denominative extensions appear after vowels (Stage 1 output) rather than directly on nominal stems and why denominal verbs can layer multiple meanings (e.g., agentive + reflexive → agent-reflexive). Languages preserve/lose/adapt at different stages (Kambaata preserves Stage 2 transparency via **-k**; Hadiyya fossilizes Stage 2 via **-t**). Stage 2 extensions (originally specific to denominal bases) generalize to apply to any verbal base.

Generalization Mechanism

The central theoretical claim is that synchronic valency-changing morphology originated as complex denominative constructions that generalized. The original function (denominal) was restricted to verbs derived from nominal bases: ***ʔ+m** → passive denominative (verbalized → passive verb) and ***y+s** → causative denominative (verbalized → causative verb). Then, the generalized function (valency) expanded to apply to any eligible verbal stem: **-am** → general passive (any transitive verb → passive) and **-is** → general causative (any root → causative).

The formal similarity between denominative markers and other verbal morphemes is due to the cliticization of the complex denominative to the nominal stem and to the generalization of some of those denominatives into

the reflexive, causative, and passive forms (e.g., **-am** ← ***ʔ-m**) due to functional similarity. The reverse pathway of verbal suffixes generalizing to nominal domains is also attested (e.g., glottalized reflexives used as denominatives).

The passive denominative (***ʔ+m** on verbalized stems) and general passives (valency alternation on any transitive verb) share a valency-reducing function. This functional overlap enabled categorial expansion; for example, once speakers recognized functional equivalence, the formal pattern extended beyond the original denominal restriction.

Therefore, numerous valency morphemes are better analyzed as generalized denominative patterns. The formal structure (verbalizer + extension) is preserved through functional expansion.

Unified System

The preceding analysis supports a unified system: support verbs grammaticalize as verbalizers, enabling nominal stems to participate in standard verbal morphology.

The denominative system and valency morphology form a continuum rather than discrete categories. The most transparent is the simple denominatives (***ʔ**, ***y** alone, §§4.2.5.1–4.2.5.2) with compositional noun → verb meaning. The intermediate is the complex denominatives (***ʔ**+extension, ***y**+extension, §4.2.5.3–4.2.5.4) with a layered structure. The most grammaticalized are the general valency markers (§§4.2.5.5–4.2.5.7) with an expanded functional domain (§§4.2.1–4.2.3).

All points on this continuum share the fundamental structure: Verbalizers enable category change from noun → verb (denominal) or are utilized within the verbal system (valency alternation).

This unified architecture contrasts with analyses treating denominatives and valency morphology as unrelated subsystems. The historical connection motivates synchronic integration.

Semantic Specialization and Lexicalization

The semantics of the denominalized stems is not always transparent. The causative denominative often adds semantic components beyond simple causation. This is most evident in cases like example [392] in §4.2.5.8, where

the relationship between derived and base forms is less predictable than in standard causatives.

Older lexicalized denominals show semantic drift. The meanings diverge from compositional predictions. Newer generalized valency uses show compositional semantics, e.g., the passive **-am** and causative **-is** add predictable meaning to verbal bases. Semantic opacity is diagnostic of historical depth, as non-transparent denominals represent earlier formations that have lexicalized, while transparent valency uses represent more recent generalizations, maintaining compositionality.

Therefore, not all formal patterns with verbalizer reflexes (**i**, **a**) have transparent denominal meanings synchronically. Morphemes (***y**, ***ʔ**) can be reconstructed from formal structure even when semantic compositionality has been lost.

The cross-linguistic variation in denominative marking within HEC reflects different historical trajectories following language divergence from PHEC. Kambaata preserves more distinct forms of the agentive **-k**/**-h** (§4.2.5.4), showing retention of ***-k** as a productive agentive, transparency in Stage 2 layering, and resistance to generalization/lenition in some patterns. Gedeo and Hadiyya show evidence of the other agentive ***-t** (§4.2.5.3), indicating the retention of a PHEC suffix. Sidaama exhibits transparent optional retention of the verbalizer **i** in a position where the other languages show analogy-driven reanalysis (§4.2.5.7) by dropping it, treating the denominative reflex as phonological (epenthetic). This collapses the denominative/valency formal distinction. Burji, Gedeo, and Sidaama show contact-influenced innovations in the reflexive denominative (§4.2.5.5), attaching the borrowed **-(V)d'** to the verbalized stems as an innovative derivation alongside inherited patterns. A separate branch status (not PNHEC) is visible in differential denominative patterns, e.g., Kambaata and Hadiyya patterns in §§4.2.5.7, 4.2.3.3.

In summary, each language preserves different stages and aspects of the original PHEC denominative system, enabling reconstruction through triangulation: Kambaata evidence for ***-k**, Gedeo/Hadiyya evidence for ***-t**, universal evidence for ***ʔ/*y** verbalizers, and differential generalization patterns documenting the timeline of grammaticalization.

The following two schemas present the reconstructed pattern and the surface reflexes attested in the HEC languages, with cross-references to subsections showing where each pattern is documented:

Verbalizer	Agentive ***-t**	Agentive ***-k**	Refl **-ʔ**	Pass ***-m**	Caus ***-s**
ʔ	**-at** (§4.2.5.3)	**-eek/-eeh** (§4.2.5.4)	**-ʔ** (§4.2.3.1)	**-am** (§4.2.1)	**-ees** (§4.2.2.2)
	-	-	**-ak'k'** (§4.2.3.2)	**-**	**-**
y	**-it** (§4.2.5.3)	-	-	-	**-is** (§4.2.2.1)
	-	-	**-i-t't'** (§4.2.3.3)	-	

The reconstructed combinations show the original verbalizer + extension sequences.

Verbalizer	Agentive ***-t**	Agentive ***-k**	Refl ***-ʔ**	Pass ***-m**	Caus ***-s**
ʔ	***ʔ-t-**	***ʔ-k-**	***ʔ-ʔ-**	***ʔ-m-**	***ʔ-i-s-/*ʔ-s-**
	-	**-**	***ʔ-k-ʔ-**	**-**	**-**
y	***y-t-**	**-**	**-**	**-**	***y-s-**
	-	**-**	***y-t-ʔ-**	**-**	**-**

The two schemas present the full combinatorial space of verbalizer (***ʔ, *y**) with extensions (***-t, *-k, *-ʔ, *-m, *-s**), showing which combinations are attested (with cross-references), which combinations are unattested (marked with "-"), how surface forms relate to proto-reconstructions, and where detailed discussions appear in the subsection of §4.2.5. The systematic gaps (unattested combinations) constrain the reconstructed system, showing that not all logically possible combinations existed. Attested patterns show that layering follows principled constraints, not free combination. **Figure *4*** summarizes the verbal derivation morphemes (§4.2) and their development, highlighting the role of the denominative constructions.

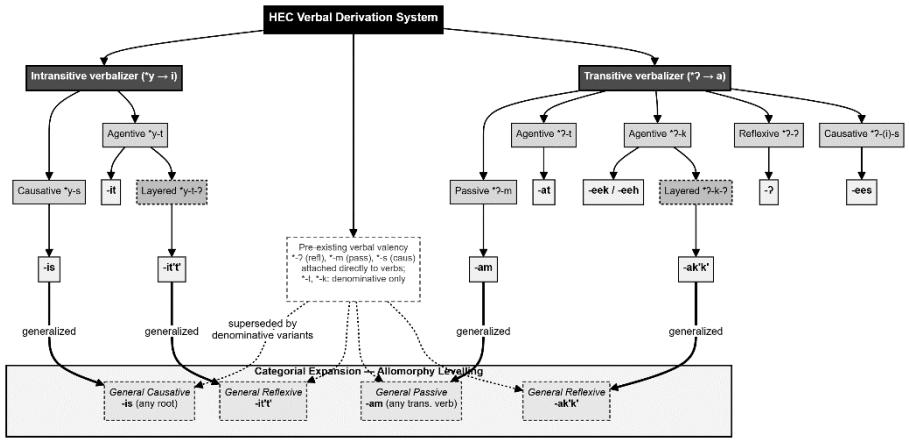

Figure 4: *The Development of the HEC Verbal Derivation System*

The preceding subsections identify the principal verbal derivational markers that can be compared across HEC and evaluate which forms are reconstructable. The comparative distributions show both continuity in the core derivational inventory and language-specific changes in particular markers and combinations. Where the evidence is uneven or restricted, reconstruction remains provisional.

The next section synthesizes the nominal and verbal reconstructions and presents a comparative summary of the morphemes discussed in this chapter. The goal is to make the distribution of reconstructed markers explicit and to highlight the most diagnostic morphological evidence for the classification developed in Chapter 5.

4.3 Summary and Conclusions

This section synthesizes the nominal and verbal reconstructions developed in this chapter and highlights the morphological evidence that bears on the subgrouping argument presented in Chapter 5. **Table 7** provides a comparative index of the nominal and verbal morphemes discussed in this chapter, and **Table 8** summarizes the reconstructed personal pronoun paradigm. The distributions support two generalizations: (a) much of the core derivational architecture is reconstructable to PHEC even when individual exponents show conditioned allomorphy or uneven attestation; and (b) a smaller set of shared innovations within NHEC provides diagnostic morphology for the Burji vs. NHEC split and for internal NHEC relationships.

Where the evidence is restricted or ambiguous, the proto-level assignment (PHEC vs. PNHEC) remains provisional.

Table 7 is intended as a comparative index to provide a compact basis for evaluating proposed PHEC reconstructions against PNHEC innovations by making comparative evidence and distribution explicit. Conventions are as follows: An asterisk marks reconstructed forms; hyphens indicate bound morphemes; and a dash indicates that a morpheme is unattested or not recoverable for the relevant language on the basis of the evidence assembled in this chapter. Where a marker shows conditioned allomorphy, the table lists the most general exponent and gives major alternants in parentheses; conditioning is discussed in the relevant subsection.

Category	PHEC	Burji	PNHEC	Hadiyya	Kambaata	Sidaama	Gedeo
Nominal							
Singulative	***-ʃ (*ʃ-ʃ, *-i-ʃʃ, *-c)**	-iccoo, -coo, -ʃo(o), -ʃaa	***-ʃ (*ʃ-ʃ, *-i-ʃʃ, *-c)**	-icco, -co	-iccu, -cu, -eeccu, ʃʃu/ccu	-icco, -co, -ʃʃo	-icco, -co, -ʃo, -jo
Reduplication plv	***C_i~C_i-a**	-	***C_i~C_i-a**	-a	C_i~C_i-a	C_i~C_i-a	C_i~C_i-a
Plurative	***-uw-a**	-	***-uw-a**	-uwwa	-	-uwa	-uwwa
Collective	***-ey-e**	-aayyaa, -aayyee	***-ey-e**	-	-	-e(e)yye	-eeye
Plurative	***-n**	-na	-	-	-	-aano	-
Deverbative	***-ʃ**	**-aancoo**, **-issoo**	***-ʃ**	-anca, -ca, -i-ʃʃa	-c-a(-ta), -iʃʃ-a(-ta)	-aan-co, -an-ʃo	-an-jo, -acco
Genitive	***ʔ**	-ng, -nd, -kk, -tt	***-ʔ**	-ʔ, -C'C'	-V́(V) (excluding u)	-ú, -í, -te, -◌́, -i, -ø	-nk, -k, -nt', -(t')t'
Masculine	***k**	k	***k**	k	k	k	k
Feminine	***t**	t	***t**	t	t	t	t
Masc. (nom)	***-u ← *-w**	-u (def)	***-u**	-i	-u (dem)	-u, -i	-i
Fem. (nom)	***-i ← *-y**	-i (def)	***-i**	-i	-i (dem)	-	-
Masc. (acc)	***-o ← *-óo**	-	***-o**	-o (TV)	-a (TV)	-o (TV)	-a (TV)
Fem. (acc)	***-e ← *-ée**	-	***-e**	-e (TV)	-e (TV)	-e (TV)	-e (TV)
Verbal							
Causative	***-s**	-s, -is	***-s**	-s, -is	-s, -is, -ʃ	-s, -is, -ʃ	-s, -is, -ʃ
Passive	***-am**	-am	***-am**	-am	-am	-am	-em
Reflexive	***-ʔ**	-ʔ, -d', -ad', -eed'	***-ʔ**	-ʔ, -i-t't', -ak'k'	-ʔ, -ak'k'	-ʔ, -id'	-ʔ, -ed'
Denominative	***ʔ(-C-C), *y(-C-C)**	-ʔ, -ad', -eed'	***ʔ(-C-C), *y(-C-C), *ʔ-k, *ʔ-t**	ʔ, -at, -it, -it't'	ʔ, -eek/-eeh, -it't'	ʔ, -id'	ʔ, -at, -ed'

Table 7: *Reconstructed HEC Morphemes and Their Reflexes*

The comparative distributions in **Table 7** show continuity in a core derivational inventory alongside language-specific changes in particular markers and combinations. Number morphology and nominal derivation exhibit both stable templates and variable exponents. The singulative suffix is reconstructable, but its surface form is conditioned: An earlier ***-ʃ** is generalized as ***-c** in many environments, with additional allomorphs determined by context. Reduplication of the final consonant recurs as a number-marking strategy in a subset of languages (Gedeo, Sidaama, Kambaata), and the plurative ***-uw-a** is best supported where it co-occurs with this reduplicative pattern (in Sidaama, Hadiyya, and Gedeo). Collective marking (***-ey-e**) is concentrated on adjectives (and a smaller set of nouns), with Burji showing distinct allomorphy. A further plurative ***-n** may underlie Burji **-na** and Sidaama **-aano**, but the comparative evidence remains limited and the reconstruction should be treated as provisional and mainly supported by fossilized evidence of ***-n** in the plural pronominals.

Inflectional morphology provides complementary evidence for both reconstruction and subgrouping. At the PHEC level, the gender consonants ***-k** (masculine) and ***-t** (feminine) are securely reconstructable given their uniform attestation across demonstratives, definiteness markers, and case morphology. Case marking is additionally inferred from terminal vowels, most transparently in pronominal and demonstrative paradigms, though substantial leveling in individual languages obscures earlier contrasts. The genitive exponent ***-ʔ** is more tentative: Direct segmental evidence is restricted, but it can be traced through conservative glottal retention and indirect reflexes (glottalization, metathesis, and accentual effects).

Verbal derivation shows a similarly template-driven organization. The denominative base is built from a verbalizer followed by an extension, and the template itself is reconstructable to PHEC even where individual complex forms are unevenly distributed. The passive is uniformly attested as ***-am** with a raised vowel in Gedeo **-em** and it is proposed to have received its vowel exponent via grammaticalization in the denominative construction. The causative suffix ***-s** is widespread, with a palatalized allomorph **-ʃ** in Kambaata, Sidaama, and Gedeo. Reflexive morphology offers one of the clearest cases for where disciplined reconstruction matters for subgrouping. A core reflexive ***-ʔ** is reconstructible to PHEC on the basis of its full HEC distribution and regular NHEC morpheme boundary alternations (including metathesis after sonorants and fusion after obstruents). In contrast, the **-(V)d'** variants found in Burji, Gedeo, and Sidaama are best treated as secondary, contact-driven changes (ultimately attributable to Oromo), and therefore

should not be projected to PHEC/PNHEC. Hadiyya and Kambaata innovated distinct reflexive exponents (**-i-t't'** and **-ak'k'**). This split between inherited suffixes, innovated structure, and a borrowed exponent underscores why diagnostic claims must rest on distribution, conditioning, and system fit rather than on isolated similarities.

The denominative analysis (§§4.2.4–4.2.5) establishes five core principles:

> Grammaticalization pathway: Support verbs [425] *ʔ 'to do' and [388] *y 'to say, do, be' grammaticalized into bound verbalizers, forming the denominative base (§4.2.4, 4.2.5.1, 4.2.5.2).
>
> Two-stage derivational template: Stage 1 combines nominal + verbalizer → verbalized stem; Stage 2 attaches extensions (agentive ***-t/*-k**, reflexive ***-ʔ**, passive ***-m**, causative ***-s**) → complex verbalized stem.
>
> Diachronic generalization: Complex denominatives (**ʔ+m** → ***-am, *y+s** → ***-is**) underwent functional expansion, being reanalyzed as general valency morphology (§4.2.5.9).
>
> Unified verbalizer system: Support verbs ***ʔ/*y** function identically across denominal (noun → verb) and valency (verb → verb) derivation.
>
> Differential grammaticalization: Language-specific variation (preservation in Burji, reanalysis in NHEC, contact-induced replacement) enables reconstruction of diachronic stages and supports subgrouping (Chapter 5).

A recurring conclusion across the verbal subsections is that diachronic change targets individual exponents more readily than it disrupts derivational templates. In NHEC, agentive extensions (***-t**, ***-k**) participate in stable template structures even as productivity shifts between competing markers and as some reflexes erode. The same template persistence is visible in the grammaticalization pathway from support verbs to verbalizers and onward to valency morphology, where complex denominative structures can fuse and be reanalyzed as synchronic passive/causative/reflexive morphology.

Table 8 presents a systematic reconstruction and a condensed comparative overview of the PHEC and PNHEC personal pronoun paradigm across three cases (nominative, accusative, possessive) and all person/number

distinctions. The paradigm illustrates several key morphological developments: (i) metathesis of possessive glottal stems, which serves as a diagnostic PNHEC innovation; (ii) paradigm leveling in the 2pl series, where glottal insertion (***ati-n-** → ***aʔ-n-**) and restructuring create analogical regularization; and (iii) stability of 3rd plural forms, which maintain PHEC shapes through both proto-stages. Detailed analysis of these reconstructions is provided in §4.1.6 (personal pronouns) and §4.1.4.2 (possessive/genitive pronouns).

Table 8: *Reconstructed HEC Pronouns: Nominative, Accusative, and Possessive*

1sg	**[2]**	**nom**	**[103]**	**acc**	**[135]**	**poss**
PHEC	***an-i**	‘I’	***e-y**	‘me’	***ʔ-y**	‘my’
PNHEC	***ani**	‘I’	***ey**	‘me’	***ʔy**	‘my’
1pl	**[63]**		**[197]**		**[188]**	
PHEC	***naʔ-**	‘we’	***n-y**	‘us’	***ni-ʔ-**	‘our’
PNHEC	***naʔ-**	‘we’	***n-y**	‘us’	***niʔ-**	‘our’
2pl	**[104]**		**[196]**		**[187]**	
PHEC	***ati-n-**	‘you’ (pl)	***ki-n-**	‘you’ (pl)	***kin-ʔ**	‘your’ (pl)
PNHEC	***aʔ-n-**	‘you’ (pl)	***kiʔn-**	‘you’ (pl)	***kin-ʔ-** → ***kiʔn-**	‘your’ (pl)
2sg	**[27]**		**[257]**		**[253]**	
PHEC	***at-i**	‘you’ (sg)	***ke**	‘you’ (sg)	***ki-ʔ-**	‘your’ (sg)
PNHEC	***at-i**	‘you’ (sg)	***ke**	‘you’ (sg)	***ki-ʔ, *-ʔ-ki**	‘your’ (sg)
3sg.f	**[6]**		**[151]**		**[366]**	
PHEC	***is-i, *is-y**	‘she’	***is-e, *is-é**	‘her’	***ise-ʔ-**	‘her’
PNHEC	***is-i**	‘she’	***is-e**	‘her’	***ise-ʔ-**	‘her’
3sg.m	**[7]**		**[292]**		**[279]**	
PHEC	***is-u, *is-w**	‘he’	***is-o, *is-ó**	‘him’	***isa-ʔ-**	‘his’
PNHEC	***is-u**	‘he’	***is-ó**	‘him’	***isi-ʔ-**	‘his’
3pl	**[11]**		**[86]**		**[318]**	
PHEC	***is-n-**	‘they’	***is-n-**	‘them’	***isn-ʔ-**	‘their’
PNHEC	***is-n-**	‘they’	***is-n-**	‘them’	***isn-ʔ-**	‘their’

The pronominal paradigm summarized in **Table *8*** supports a compact set of PHEC reconstructions that integrate case, number, and possession. Third-person forms share a stem **is-* across nominative and accusative, with case-marking for feminine (***-i** nominative vs. ***-e** accusative) and for masculine (***-u** nominative vs. ***-o** accusative). In the plural, a segmentable plurative ***-n** is recoverable in several forms, while possessive morphology allows a genitive ***-ʔ** to be inferred even where direct segmental reflexes are reduced.

The morphological reconstruction demonstrates that HEC synchronic diversity arises from three interacting processes: (a) regular sound change applying consistently across morpheme boundaries (Ch. 3); (b) analogical paradigm

leveling creating language-specific irregularities; and (c) grammaticalization pathways differentially advanced across daughter languages. The reconstructed morphemes summarized in **Table 7** and **Table 8** provide diagnostic morphology for the HEC subgrouping developed in Chapter 5. Burji reflects archaic retention in several domains, while NHEC exhibits shared innovations such as including metathesis in the pronominal system, causative palatalization, and changes in agentive extensions, which define the northern branch. Across nominal and verbal morphology, the same grammaticalization trajectory recurs (support verb → verbalizer + extension → valency marker), consistent with a template-based derivational structure inherited from PHEC. Future work should target the weakest points in the current reconstruction – i.e., especially genitive alignment and case morphology – to test whether they align with the same branch-defining innovations or reflect independent innovations and contact.