

Article

Preliminary Notes on Kwegu*

Shuichiro NAKAO
(The University of Osaka)
s.nakao.hmt@osaka-u.ac.jp

Abstract

Kwegu is a less studied Southeastern Surmic language spoken in the South Omo zone in southern Ethiopia. This is a preliminary report of recent fieldwork conducted by the author in March 2025 in Jinka. The data in this study will shed light on some hitherto unknown lectal variations and grammatical aspects of Kwegu, including the phonological statuses of glottal sounds and tone, the basic nominal and verbal morphosyntax, and information structure.

1 Introduction

Kwegu (ISO639-3: xwg) is a less studied Southeastern Surmic (Nilo-Saharan) language spoken in the South Omo zone in southwestern Ethiopia. This language has been studied mainly by two Japanese linguists, Osamu Hieda (Hieda 1990, 1991, 1992, 1998, 2024) on the Kuchuru dialect and Hideyuki Inui (2011a, 2011b, 2013a, 2013b, 2014) on the Galgida dialect, who have provided robust data on its lexicon and grammar showing dialectal contrasts. These studies revealed Kwegu's typological peculiarities not attested or rare in the rest of the Surmic branch, such as the five vowel system, phonological ejectives (also attested in Me'en) and lack of the complex nominal morphology (e.g., case-related endings).

This language is variously spelled (apart from its exonyms, such as Muguji, Yidinich, etc.). While it is listed as *Kwegu* in major databases including ISO639-3 and Glottolog, it is referred to as *Koegu* by Hieda and *Kowegu* by Inui, while my informant¹

* This research was supported by KAKENHI Grant-in-Aid (Grant Number: 22401046). I appreciate Mr. A who provided me with Kwegu data, Dr. Firew Girma Worku for arranging the fieldwork, and two anonymous reviewers for their helpful comments.

¹ My informant is Mr. A from the Kuchuru village (20 years old in March 2025). The research was conducted in Amharic by me. According to him, he passively understands Nyangatom and Kara.

pronounces [kòjgù] and writes in the Ethiopic script as ኮጃጉ /kojgu/². However, this report adopts the currently prevalent form, *Kwegu*, for convenience.

This study is not meant to be a balanced grammatical sketch of Kwegu. It aims at providing some fresh data on the hitherto underdescribed lectal variances and structures of this language, based on my own fieldwork in Jinka in March, 2025. The data are presented in IPA, except that long vowels are transcribed with doubled vowels (**aa** instead of **a:**).³ Data from the previous studies are cited in this system.

2 Phonology

This section provides some information on phonological aspects of Kwegu. See Hieda (1992, 2024; abbreviated H92, H24) and Inui (2014; abbreviated I14) for the proposed full segmental inventories of Kwegu. In this section, the phonetic realization is written in brackets [...] and the phonemic analysis in slashes /.../ (which are basically omitted in Sections 3–8). In the phonemic transcription, I use floating tone ` (see 2.2).

2.1 Glottal sounds

Although both H92/H24 and I14⁴ list /h/ as a Kwegu phoneme, they differ in its usage. My data attest to more cases of word-initial /h/ (with phonetic variations) that correspond to what Hieda and Inui describe as onset-less. This may suggest that it is perhaps a phonetic realization of word-initial onset-less syllables.⁵

(1)	H24	I14	my data	
	a	a	[há ~ fhá ~ á]	copula, ergative marker
	a-	(h)aʔ-	[hà- ~ fhà- ~ à-]	first person verbal prefix
	i-	jeʔ-	[hì- ~ fhì- ~ jì ~ ì-]	second person verbal prefix
	hiʔiŋ	xin ~ -ŋ	[hiʔŋ ~ fiʔŋ ~ jiʔŋ ~ iʔŋ]	proximate demonstrative
	hiŋafi	iŋafi	[híŋàf ~ fíŋàf ~ jíŋàf ~ íŋàf]	‘girl’
	hur ~ ur	ur	[húr ~ fúr ~ wúr ~ úr]	‘person’

² More precisely, **tók ká kòjgù** (mouth GEN Kwegu) ‘the Kwegu language’.

³ The realization of vowels is reduced to the 5-vowel system following previous studies, although more detailed studies are desiderata. Hieda basically does not distinguish short and long vowels, but Inui does. My data seem to support Inui’s.

⁴ I14 lists /ɣ/ as a phoneme which corresponds to either /k/ or /h/ of Hieda (1992). My data also attested to an instance of [x] as a postvocal realization of the genitive particle **ká** (**á ká [xá] kòjgù** (COP GEN Kwegu) ‘It is Kwegu’s [language]’).

⁵ For instance, in an utterance, my informant pronounced Amharic **af** አፍ with [h], i.e., [háf]. In Nyangatom, too, word-initial onset-less syllables tend to have [h] ~ [f], e.g., **(h/f)ájǝŋ** ‘1SG’ (based on my own data, collected in Jinka in March 2025).

However, there are also some instances that consistently either lack or have /h/ in all H92/24, I14 and my data.

(2)	H92/24	I14	my data	
	an	aan	[áân] (not *háân)	‘1SG’
	in	iin	[íîn] (not *híîn)	‘2SG’
	arte	arte	[àrtè] (not *hàrtè)	‘fish’
	hánta	hanta	[hántà]	‘child’
	áhur	ahur	[àhùr]	‘four’

The status of [h] in non-word-initial positions is rare and less clear. My data attest to a few cases of non-word-initial [h] that differ from H92 and/or I14.

(3)	H92	I14	my data	
	ḍaa	ḍaxa	[ḍâhá] ~ [ḍâḥá]	‘two’
	la	lax	[láh]	‘six’
	ts’oba	tsoba	[tʃ’óbâh]	‘seven’

H24 is dubious about the status of [ʔ] as a phoneme, while I14 lists it as a phoneme. Here again, my data show some difference from the previous studies. My data in (5) can be interpreted as the loss of etymological [ʔ].

(4)	H24	I14	my data	
	toʔo	toʔo	[tóʔò]	‘house’
	hiʔiŋ	xiŋ	[(h)iʔŋ]	‘PROX.SG’
	gitaʔan	gitaŋ	[gítàŋ]	‘PROX.PL’
	a-am-ba	(h)aʔ-aam-ba	[(h)à(ʔ)áâmbà]	‘I eat’ (1SG-eat-IPFV)
	-aa	-(j)aʔa	[-à(ʔ)à]	3SG perfective verbal suffix

(5)	H24:	hiʔiŋ ka ham (PROX.SG GEN there)
	I14:	hiʔiŋ ka ham ~ iŋkaam
	my data:	[iŋkáàm] ~ [ḡkáàm] (but hàʔm ‘there’)
		‘that (DIST.SG)’

At this point, it is still early to judge the phonological status of [h/h] and [ʔ]. Note that previous studies and my data also show differences about the realization of the other sounds (e.g., vowel length, ejective/implosive), but their relationship is not clear enough to be argued in this study.

2.2 Tone

Kwegu seems to have phonemic distinctions of two level tones, high and low, but there are some notable realizations of these tonemes.

Some lexemes have falling tones in certain contexts. It can be analyzed as a floating tone /ʔ/, whose realization is optional when it is followed by a low-toned syllable.

- (6) a. [áân] ‘I’
 /áân`/
 1 SG
- b. [áân àmâtbà] ~ [áán àmátbà] ‘I drink’
 /áân` à-mát-`=bà/
 1 SG 1-drink-SG=IPFV

When it is followed by an underlyingly high-toned syllable, the floating tone is realized as the low or mid tone on that (i.e., a type of downstep). The floating tones, at least partially, seem to be derived from deletion of a vowel. Note that the verbal form **mát-`** ‘drinks’ in (7a,b) corresponds to **mati** in H24.

- (7) a. [ífi mát òò (ōō)] ‘What does he drink?’
 /ífi mát-` óó?/
 3 SG drink-(3)SG what
- b. [úr i?η mát òjǔ (ōjǔ)] ‘This person drinks coffee (as an answer to (7a)).’
 /úr i?η mát-` ójǔ./
 person PROX.SG drink-(3)SG coffee
- c. [íngîn màtbà ójǔ] ‘They drink coffee.’
 /íngîn `-mát=bà ójǔ./
 3 PL 3 PL-drink=IPFV coffee

Words without a high tone may realized with word-initial mid tone in the isolation form. Compare the realization of this mid tone (8) with the high tone (9).

(8)	/iʔŋ/	[iʔŋ] ~ [iʔŋ̃]	‘this’	[úr iʔŋ]	‘this man’
	/àrtè/	[ārtè]	‘fish’	[àrtè gítàŋ]	‘these fishes’
(9)	/sél’/	[sêl]	‘9’	[tàbí sêl]	‘19’
	/tóʔò/	[tóʔò]	‘house’	[tóʔò gítàŋ]	‘(these) houses’

3 Personal pronouns

Personal pronouns have an independent form and two genitive forms as summarized in Table 1 (cf. H92: **irun**, **ifi** ‘3SG’, **galgitaʔan**, **galgitaiʔi** ‘3PL’; I14: **iŋan** ‘2SG.F’, **iŋ** ‘3SG.M’, **iŋki** ‘3SG.F’). Neither H24 or I14 gives the genitive 2 forms. I14 distinguishes grammatical gender in the personal pronouns and the verbal paradigms (which is unique in Surmic), but they are not attested in my data.

Table 1. Personal pronouns

	independent	genitive 1	genitive 2
1SG	áán`	nàànì	-ʔín-nààn-ù
2SG	íín`	nùùnù	-ʔín-nùùn-ù
3SG	ífi	nììnì	-ʔín-nììn-ù
1PL	úwòù	gùàjù	-ʔíŋ-gùàj-ù
2PL	íjòù	gùùjù	-ʔíŋ-gùùj-ù
3PL	íŋgín`	gìèjì	-ʔíŋ-gìèj-ù

The independent forms are used as both subject and object. Kwegu seems to require overt pronouns even when they are indicated in the verbal inflection.

- (10) a. áán` à-náǵĩ-` íín`=bà. ‘I love you.’
 1SG 1-love-SG 2SG=IPFV
 b. íín` ì-náǵĩ-` áán`=bà. ‘You love me.’
 2SG 2-love-SG 1SG=IPFV

The genitive-1 form has both modificational and nominal uses, while the genitive-2 form has only modificational use. At this point, the difference in (11) is unclear. The analysis of the genitive-2 form remains a problem, which is reminiscent of the Surmic genitive construction. Since none of the previous studies attest it, more data are needed to discuss if it represents a retention or a new development.

- (11) **tóʔò nààni** ‘my house’
tóʔò-ʔín-nààn-ù ‘my house’

- (12) **tóʔò iʔŋ á nààni.** ‘This house is mine.’
house PROX.SG COP GEN.1SG

On the other hand, nouns do not have any case-based inflections (including tonal alternation) as already indicated by H92. The genitive construction is coded with the particle **ká** ‘GEN’, which can also be the head of the noun phrase. The genitive particle can be omitted in compounding such as (14).

- (13) **iʔŋ á tóʔò ká síndí.** ‘This is Sindi’s house.’
PROX.SG COP house GEN Sindi
tóʔò iʔŋ á ká síndí. ‘This house is Sindi’s.’
house PROX.SG COP GEN Sindi

- (14) **tóʔò ká tàrààrí ~ tóʔò tàrààrí** ‘school’
house GEN student house student

4 Demonstratives and nominal morphosyntax

My data attest to two-way (proximal vs. distal) distinction of demonstratives (with corresponding adverbs, **óʔúŋ** ‘here’ and **hàʔm** ‘there’). The third person pronouns given by H24 and I14 seem to be demonstratives, according to my informant. H24 gives the medial series (SG **haʔaŋ**, PL **gitaʔan haʔaŋ**), but my informant did not know these forms (as with I14). Proximal demonstratives are often used by my informant to translate an Amharic definite article.

Table 2. Demonstratives

	SG	PL
PROX	ìṛṇ	gítàṇ
DIST	ìṇkáàm	gítàj, gítàṇkáàm

Demonstratives are postposed to a noun in their modification use (15). Some nouns attest to specific proximate forms (16).

- (15) **tóṛò ìṛṇ** ‘this house’
tóṛò gítàṇ ‘these houses’
tóṛò ìṇkáàm ‘that house’
tóṛò gítàṇkáàm ‘those houses’
- (16) **íṇáf` ìṛṇ ~ íṇáf` ìṛṇ ~ íṇáf-ṛṇ`** ‘this girl’ (**íṇáf`** ‘girl’)
àrt-óṛúṇ (*àrtè ìṛṇ) ‘this fish’ (**àrtè** ‘fish’, cf. **óṛúṇ** ‘here’)

As already indicated by H92/24 and I14, Kwegu nouns have plural marking **-Vn**, which is dropped when the number is marked by another modifier.

- (17) SG PL
úr ‘man’ **ùr-ún** ‘men’ (cf. H24/I14 **ur-an**)
úr tàṭṭí ‘big man’ **úr tàtàṭṭí** ‘big men’
(*ùr-ún tàtàṭṭí)
úr tàṭṭí ìṛṇ ‘this big man’ **úr tàtàṭṭí gítàṇ** ‘these big men’
(*úr tàṭṭí gítàṇ)

5 Numerals

As already indicated by previous studies, the numeral system of Kwegu massively consists of Cushitic/Omotc loanwords (a tendency common in Nilo-Saharan languages of the East Rift area), except for ‘1’ to ‘3’. As H92 and I14 identified, ‘6’ and above are directly from Kara (South Omotic), but ‘4’ and ‘5’ are from East Cushitic (perhaps Arbore or Dhaasanach). Note that the other Southeast Surmic languages have Cushitic/Omotc forms for ‘2’ and ‘3’ but not for ‘4’ and ‘5’.

(18)	kièn	‘1’	tàbí kàlà	‘11’	tàbí	‘10’
	ḏàhá	‘2’	tàbí làmá	‘12’	làmatàm	‘20’
	ḏièn(i)	‘3’	tàbí màkkám	‘13’	màkkámtàm	‘30’
	àhùr	‘4’	tàbí òjdí	‘14’	òítàm	‘40’
	ṭôn	‘5’	tàbí dônḡ	‘15’	dónṭàm	‘50’
	láh	‘6’	tàbí láh	‘16’	látàm	‘60’
	ṭʷóbâh	‘7’	tàbí ṭʷóbâh	‘17’	ṭʷóbátàm	‘70’
	lónkâj	‘8’	tàbí lónkâj	‘18’	ìítàm	‘80’
	sêl	‘9’	tàbí sêl	‘19’	nàʔátàm	‘90’

(19)	làmatàm kàlà	‘21’
	làmatàm làmá	‘22’
	làmatàm màkkám	‘23’

Note that, in my data, the lexeme for ‘10’ is replaced by a Kara form. While H92 and I14 give **tomon/tomun** ‘10’, this lexeme was judged as a Nyangatom word (as actually it is, and ultimately from East Cushitic) by my informant. Also, my data show unique forms for ‘80’ (cf. H92 **lonkaitam**, I14 **lonkaitam**) and ‘90’ (H92 **saltam**, I14 **selltam**), whose origins are unknown, along with Amharic loanwords for ‘100’ (**መቶ** cf. H92 **dip**, I14 **ḏif** ‘100’, Kara loanword) and ‘1000’ (**ሺህ**) as shown below.

(20)	màtó kièn	‘100’
	màtó ḏàhá	‘200’
	jí kièn	‘1000’

6 Verbal morphology

This section provides some data on verbal inflectional and derivational morphology.

6.1 Subject person marking and TAM

Kwegu has two basic series of subject person/number inflection on the verb, labelled ‘unmarked’ and ‘perfective’ forms by H92. The realization of the 3PL form differs according to whether the verb has a stem-initial consonant (i.e., **g** → `/_C). Note that it also triggers tonal alternation.⁶

⁶ H24 gives the 3PL prefix **g** ~ ‘, namely with high tone. However, in my data, the word-initial

Table 3. Verbal paradigms

	‘eat’		‘drink’	
	unmarked	perfective	unmarked	perfective
1SG	à-áám-`	à-áám-ì-jà	à-mát-`	à-mát-ì-jà
2SG	ì-áám-`	ì-áám-ì-jà	ì-mát-`	ì-mát-ì-jà
3SG	áám-`	áám-à?à	mát-`	mát-à?à
1PL	à-áám-à	à-áám-à-jà	à-mát-à	à-mát-à-jà
2PL	ì-áám-à	ì-áám-à-jà	ì-mát-à	ì-mát-à-jà
3PL	g-ààm	g-ààm-à?à	màt	màt-à?à
	(g`-áám)	(g`-áám-à?à)	(`-mát)	(`-mát-à?à)

Both H92 and I14 give the third paradigm of ‘imperfective’,⁷ which is a simple combination of the unmarked form and the suffix **-ba**. However, I tentatively consider it to be a clitic as it is not always realized as a verbal ending (see (10)).

The imperfective and perfective verbal forms are metalinguistically explained as corresponding to the imperfective and perfective forms of Amharic. Although it requires more investigation to identify the exact meaning of each form, I tentatively understand it as an aspectual one.

- (21) a. áán` à(?)-áám-`=bà dààbó.
 1SG 1-eat-SG=IPFV bread
 ‘I eat bread’ (Amh. እኔ ዳቦ እበላለሁ።)
- b. áán` à(?)-áám-ìjà dààbó.
 1SG 1-eat-1SG.PFV bread
 ‘I ate bread.’ (Amh. እኔ ዳቦ በላለሁ።)

As for the unmarked form (which I14 labels the ‘shortened form’), previous studies consider it as an aspect-neutral form.

My data attested to two hitherto unreported particles, **lémè** (probably stative past)⁸ and **ká=** (probably irrealis or future, as it is used to translate Amharic optative and

syllable of the 3PL is clearly lower than that of the 3SG form. According to I14, the Galgida dialect completely lacks the realization of /g/.

⁷ H24 switches the labels ‘perfective’ and ‘imperfective’, taking lexical aspect into account.

⁸ Perhaps this form is comparable with I14 **lemen** ‘in front’.

imperfective forms, e.g., **yəbāla** ‘he should eat’, **yəbālal** ‘he eats/will eat’). **ká=** is used with the unmarked form. It can cooccur with the imperfective **=bà**.

- (22) a. **tóʔò áné=bà.** ‘There is a house.’
house exist=IPFV
b. **tóʔò áné=bà lémè.** ‘There was a house.’
house exist=IPFV PAST

- (23) a. **úr iʔŋ ká=áám-` dààbó.**
person PROX/DEF IRR=eat-(3)SG bread
‘This/the person will eat (eats) bread.’
b. **íŋgín` ká=g`-áám=bà [kágààmbà] dààbó.**
3PL IRR=3PL-eat=IPFV bread
‘They will eat bread.’
c. **bàrjù ká=à-jím-`=bà.**
God IRR=1-praise-SG=IPFV
‘Thank you (Let me praise God).’ (Amh. እግዚአብሔር ይመስገን።)

6.2 Voice

I14 describes two multifunctional suffixes for voice alternation, **-en** (anticausative, passive, inchoative, reciprocal) and **-ifi/-ife** (causative). The former corresponds to what H24 labels intransitive marker **-Vne**, which marks intransitive (antipassive and anticausative) and passive. My data attest to **-iŋ** (perhaps cognate with **-en/-Vne**) as an antipassive marker (24) and **-ife** as an applicative marker (25).

- (24) a. **ífi kóh-`=bà rúz`.** ‘He cultivates rice.’
3SG cultivate-(3)SG=IPFV rice
b. **ífi kóh-iŋ=bà.** ‘He cultivates.’
3SG cultivate-AP=IPFV
(25) a. **ífi kù-ife síndí** ‘He went to Sindi.’
3SG go-APPL Sindi
b. **ífi dàwl-ife-jàʔà síndí.** ‘He called Sindi.’
3SG call-APPL-3SG:PFV Sindi

7 Word order and verbal morphology

The difference between the “unmarked” verb form and the imperfective (**-ba**) form has remained less investigated. While H24 and I14 consider their difference is at least partially an aspectual one, my data show that the unmarked form also has a function as a (TAM-neutral) ‘argument-focus’ form.

Also H24 notes that, although Kwegu basically has the AVO word order, there are variations such as VAO based on the information structure and this affects the verb form. My data support this idea.

7.1 Default word order

In my data,thetic sentences take the AVO order (26). This word order is the default strategy to translate Amharic AOV (non-cleft) sentences.

- (26) Q. **á** **óó** **gáán?** ‘What happened? (lit. What is it now?)’
 COP what now
- A1. **áli** **kát-à?à** **úwà.** ‘Ali hit a child.’
 Ali hit-3(SG).PFV child
- A2. **úwà** **ì?η** **kát-à?à** **áli.** ‘The child hit Ali.’
 child DEF.SG hit-3(SG).PFV Ali

7.2 Argument-focus

Wh-questions in Kwegu require specific structures, where the interrogatives basically follow the unmarked verb form (and the TAM is neutralized).

Object questions and answers always take the AVO word order with the unmarked form. Amharic object-focusing cleft sentences are translated with this structure.

- (27) Q. **áli** **kát-`** **nín?** ‘Who did Ali hit?’
 Ali hit-(3)SG who
- A. **áli** **kát-`** **úwà.** ‘Ali hit a child.’
 Ali hit-(3)SG child

- (28) **íín`**      **ì-mát-`** **óíjú.**
 2SG 2-drink-SG coffee
 ‘It is coffee that you drink/drank.’ (Amh. ቡና ነው የምትጠጣው/የጠጣኸው።)

On the other hand, a subject question is coded with the VOA or AVO word order with unmarked verb, and the subject follows what I tentatively analyze as the ergative marker **á** (~ **há**), which is phonetically identical with the copula. Note that it is not realized in non-focal sentences (26), as such it is rather a ‘focalized ergative’ marker.

Although the semantic difference between the VOA and AVO orders is unclear at this point, it shows that the pre-verbal slot can host the focalized constituent.

(29) Q. **kát-`** **úwà** **á** **nín?** ‘Who hit the child?’

hit-(3)SG child ERG who

A. **kát-`** **úwà** **á** **áli.** ‘Ali hit the child.’

hit-(3)SG child ERG Ali

(30) **á** **nín** **ká=áám-`** **indžèrá?** ‘Who will eat injera?’⁹

ERG who IRR=eat-(3)SG injera

(31) **á** **àstàmààrí** **kát-`** **tàmààrí.** ‘It is a teacher who hit a child.’

ERG teacher hit-(3)SG student

At this point, it is not certain that it is truly an ergative marker since I do not have enough data on intransitive clauses. However, (32b) may suggest that focalized (or non-topic) intransitive subjects do not have it.

(32) a. **tóʔò** **áné=bà.** ‘There is a house/A house exists.’ (Amh. ቤት አለ።)

house exist=IPFV

b. **óʔúŋ** **áné** **tóʔò.** ‘Here is a house.’ (Amh. እዚህ ቤት አለ።)

here exist house

If the focalized subject is the personal pronoun, the verbal form lacks the personal prefix (33b). H24 records the similar alternation without the ergative marker.

⁹ Note that this example has the irrealis marker **ká=** with the unmarked verb form. This may imply that **ká=** is simply an adverb out of the scope of verbal morphology.

- (33) a. **á íín` ì-mát-` óííú.** ‘It is you who drink/drank coffee.’
 ERG 2SG 2-drink-SG coffee
 b. **mát-` óííú á íín`.** ‘It is you who drink/drank coffee.’
 drink-SG coffee ERG 2SG

The ergative marker may represent a grammaticalization from a copula¹⁰ (in a cleft-like structure); it cannot be analyzed synchronically as a copula anymore since it does not mark another focalized constituent, such as the object (27, 28) or the adjunct (34). As H24 states, the sentence-final position seems to be the focus slot.

- (34) **àstàmààrí iṛṇ kát-` tàmààrí iṛṇ (*á) tóḷò ká tàmààrí kièni.**
 teacher DEF.SG hit-(3)SG student DEF.SG (ERG) house GEN student LOC
 ‘It was at the school that the teacher hit the student.’
 (Amh. በትምህርት ቤት ነው አስተማሪው ተማሪውን የመታው።)

7.3 OVA word order

Kwegu also allows OVA word order, which requires a special “OVA” marker phonetically realizing as if a verbal ending. The subject/agent interrogative pronoun that occurs in this construction (superficially) lacks the ergative marker (35).

- (35) a. **ífi kát-à?à úwà.** ‘He hit a child.’
 3SG hit-PFV child
 b. **ífi [kát-à] úwà.** ‘A child hit him (It is a child who hit him).’
 3SG hit-OVA child

- (36) **áli [kát-à] nín?**
 Ali hit-OVA who
 ‘Who hit Ali? (As for Ali, who hit him?)’ (Amh. አሊ ማን ነው የመታው?)

¹⁰ The assumed directionality of this language change (COPULA > ERGATIVE) is based on the fact that, while other Southeastern Surmic languages such as Mursi (Firew 2021) and Suri (Bryant 2024) have a cognate copula, no Surmic language possesses a comparable ergative or nominative preposition.

Perhaps the ‘OVA’ suffix can be analyzed as a composite of the unmarked verb form and the ergative marker as shown below.

- (35)’ b. *ífi kát-` á úwà.*
3SG hit-(3)SG ERG child

If this analysis is correct, we can conclude that the focalized or non-topic subject always requires the ergative marker and the unmarked (or ‘focus’) verb form. This may also suggest that the “marked” verb form may be correlated with predicate focus, but further investigation is needed.

8 Conclusion

This study has provided some fresh data on some aspects of Kwegu for future discussions, such as the phonological status of [h] and [ʔ], realization of tone, lectal variances in some closed classes (personal pronouns, demonstratives and numerals), correlation of word order, verbal morphosyntax and information structure.

This study has tentatively proposed the existence of floating low tone (Section 2.2) and an ergative(-like) marking that realizes only on the non-topic agent nouns (while the verbs have nominative-type inflection). Although the data presented in this study are not enough to draw a conclusion, Kwegu may represent a unique Surmic language that has a prepositional (and functionally marked) ergative marker developed from the copula.

Abbreviations

*: ungrammatical/unacceptable, 1/2/3: person, Amh: Amharic, AP: antipassive, APPL: applicative, COP: copula, DEF: definite, DIST: distal, ERG: ergative, GEN: genitive, H24: Hieda (2024), H92: Hieda (1992), I14: Inui (2014), IPFV: imperfective, IRR: irrealis, LOC: locative, PAST: past, PFV: perfective, PL: plural, PROX: proximal, SG: singular, TAM: tense-aspect-modality.

References

- Bryant, Michael (2024) Suri (Tirmaga-Chai). In: Meyer, Ronny, Bedilu Wakjira & Zelealem Leyew (eds.) *The Oxford Handbook of Ethiopian Languages*, 834–866. Oxford: Oxford University Press.
- Firew Girma Worku (2021) A grammar of Mursi: A Nilo-Saharan language of Ethiopia. Leiden: Brill.
- Hieda, Osamu (1990) Koegu, a preliminary report. *Journal of Swahili and African Studies* 1: 97–108.
- Hieda, Osamu (1991) Koegu vocabulary, with a reference to Kara. *African Study Monographs*, Suppl. 14: 1–70.
- Hieda, Osamu (1992) A grammatical sketch of the Koegu language. *Journal of Swahili and African Studies* 3: 131–155.
- Hieda, Osamu (1998) A sketch of Koegu grammar: Towards reconstructing Proto-Southeastern Surmic. In: Dimmendaal, Gerrit J. & L. Marco (eds.) *Surmic Languages and Cultures*, 345–373. Cologne: Rüdiger Köppe.
- Hieda, Osamu (2024) Koegu (Muguji). In: Meyer, Ronny, Bedilu Wakjira and Zelealem Leyew (eds.) *The Oxford Handbook of Ethiopian Languages*, 742–757. Oxford: Oxford University Press.
- Inui, Hideyuki (2011a) Preliminary report of the research on Kowegu (in Japanese). In: Inui, Hideyuki (ed.) *Investigations and researches on the Omotic-Cushitic minority languages and the database construction using the Geographic Information System (Cushitic-Omotic Studies 2010)*, 69–89. Yamaguchi: Yamaguchi University.
- Inui, Hideyuki (2011b) Phonemic inventory of Kowegu: The Galgida dialect (in Japanese). *Journal of General Linguistics* (University of Tsukuba) 14: 41–64.
- Inui, Hideyuki (2013a) Kowegu nouns 500 (in Japanese). *Studies in Ethiopian Languages* 2: 9–28.
- Inui, Hideyuki (2013b) Kowegu adjectives (in Japanese). *Studies in Ethiopian Languages* 2: 29–38.
- Inui, Hideyuki (2014) A grammatical sketch of Kowegu (in Japanese). *Studies in Ethiopian Languages* 3: 24–49.