AFRIKA UND ÜBERSEE

VOLUME 98 Published on September 5, 2025

Trilingual Journal of African Languages and Cultures

Revue trilingue des langues et cultures africaines

Dreisprachige Zeitschrift für afrikanische Sprachen und Kulturen

Edited by the Abteilung für Afrikanistik und Äthiopistik
at Universität Hamburg

CONTENT

AFRIKA UND ÜBERSEE

VOLUME 98 Published on September 5, 2025

Trilingual Journal of African Languages and Cultures

Revue trilingue des langues et cultures africaines

Dreisprachige Zeitschrift für afrikanische Sprachen und Kulturen

Edited by the Abteilung für Afrikanistik und Äthiopistik
at Universität Hamburg

Imprint

Editing institution
Abteilung für Afrikanistik und Äthiopistik at Universität Hamburg

All correspondence concerning editorial matters should be sent to

Universität Hamburg
Asien-Afrika-Institut
Abt. für Afrikanistik und Äthiopistik
Edmund-Siemers-Allee 1 / Flügel Ost
20146 Hamburg
Germany

E-Mail: afrikaunduebersee@uni-hamburg.de

Volume DOI 10.15460/auue.2025.98.1
Journal DOI 10.15460/auue
ISSN 0002-0427
eISSN 2749-0971

Technical hosting
Hamburg University Press
Publishing house of the Hamburg State and University Library
(Staats- und Universitätsbibliothek Hamburg)
Von-Melle-Park 3
20146 Hamburg
Germany

Published with the support of
Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)

CONTENT

STAMP morphs, prefix conjugations and multiverb predicates in the Chadic languages

Gregory D. S. Anderson

Living Tongues Institute for Endangered Languages & University of South Africa
gdsa@livingtongues.org

Abstract
STAMP morphs, prefix conjugations and the multiverb predicate subtype called auxiliary verb constructions (AVCs) that these derive from in the Chadic languages are discussed. Chadic languages are typically SVO (or VSO) and typically AUX V syntactically. Synchronically complex verb forms which derived etymologically from AVCs are also discussed. Across all subgroups of Chadic (but most common in West and Central), subject pronouns have sometimes fused with auxiliaries that have eroded to only a tone or floating tone or have become completely opaque synchronically. In recent Africanist typological studies these have been called STAMP morphs. These occur in different inflectional configurations across the languages of the family speaking to their origins as AVCs. There are at least ten STAMP morph series in Guus (Sigidi) (Caron 2002). In some Chadic languages like Mbuko, STAMP morphs have later fused with lexical verbs to create prefixal conjugations. This process of deriving prefix conjugations from fused STAMP morph constructions has happened repeatedly in the history of Chadic. This is a major reason why Chadic prefixal conjugations are not directly inherited from Proto-Afroasiatic.

Keywords: Chadic, auxiliary verb constructions, STAMP morphs, prefixation, historical linguistics

1 Introduction

In this paper I discuss the typology of a multiverb predicate subtype, the auxiliary verb construction (AVC), focusing on data from the Chadic language family, offering data from forty-six different Chadic languages spanning the full taxonomic diversity of the family from almost all the subgroups of West, Central, Masa and East Chadic. I follow the pan-chronic, constructional-functional typology

DOI 10.15460/auue.2025.98.1.394

of Anderson (2006, 2011) that views auxiliary verb constructions as a specific type of semantically determined multi-verb construction.[1]

In general terms, in two-part multi-verb predicates one speaks of a $verb_1$ and a $verb_2$. If one of the verbal elements is lexical and the other one expresses a grammatical or functional modification of the lexical verb, for example grounding it in a larger discourse space (e.g., encoding tense, aspectual or modal modification, polarity, arguments/participants, etc.), then the structure is an AVC.

AVCs are thus here understood as mono-clausal multi-verb phrases that minimally consist of an auxiliary verb component that contributes some grammatical content to the expression and a lexical verb component that contributes lexical content to the expression. An 'auxiliary verb' is thus an element on the full-lexical-verb-to-bound-functional-element grammaticalization chain (ex. 1) that performs some more or less definable grammatical function.

(1) $\textit{lexical verb}_{\text{[+ syntagma]}} > \textit{auxiliary verb}_{\text{[+ lexical verb]}} > \textit{affix}_{\text{[-verb.head-]}} > \emptyset$

This well-known path collapses in(ter)dependent functional and phonological/morphotactic hierarchies and exemplifies the inherently diachronic nature of AVCs, whereby one component element (the AV) changes semantically from a LV into an AV, loses morphophonological characteristics of an independent word and ultimately becomes an affix. Subsequently such elements may erode to zero and a new grammaticalization cycle might begin; see discussions in Bybee & Dahl (1989), Heine (1993), Bybee et al. (1994) or Kuteva (2001).

In section 2 we outline the cross-linguistic inflectional typology of AVCs and exemplify these using data from Chadic languages in section 3. In sections 4 and 5 we examine two kinds of historical processes involving AVCs. One is the fusing of the LV and the AV into morphologically complex words examined in section 4. In section 5 we present data on forms that represent the reinterpretation of

1 The analysis presented here is constructional as it entails the interaction of two component elements, the auxiliary verb (AV) and lexical verb (LV) each contributing to the meaning of the resulting multi-verb predicate construction. It is functional as the AV definitionally contributes functional or grammatical meaning to the whole. It is pan-chronic insofar as it deals with synchronic multi-verb predicates, as well as synchronically complex verb forms that are morphotactically or morphophonologically single words but which derive etymologically from AVCs, also sometimes via intermediate stages involving STAMP morphs.

DOI 10.15460/auue.2025.98.1.394

often highly eroded auxiliaries with subject marking as subject pronouns themselves that include TAM specifications. These elements have come to be called STAMP morphs in the Africanist typological literature (e.g., Anderson 2012, 2015, 2017). Finally, in section 6 I examine the rise of prefixal conjugation markers that resulted from the fusing of STAMP morphs and LVs that they formed constructions with to create single complex verbal words, a process which has happened several times in the history of different Chadic subgroups. This explains why such prefixal conjugational elements are not necessarily cognate across different Chadic subgroups and importantly also why Chadic prefixal inflectional series do not appear to be cognate with those attested in other Afroasiatic branches such as Semitic.

2 Inflectional typology of AVCs in Chadic[2]

Among the most meaningful ways to typologize auxiliary verb constructions include what their diachronic source constructions are or their subsequent historical developments. But perhaps most revealing is the patterning of the indexation of the functional (inflectional) features required by the grammar. The varied patterns of inflection in AVCs generally reveal their syntactic source construction and often the lexical origins of the auxiliaries that have been grammaticalized.

The concept of headedness in AVCs can be approached in several different ways, e.g., semantically, syntactically, functionally (i.e., inflectionally or morphosyntactically), and morphophonologically or prosodically. Semantically speaking, in AVCs, one component is designated as the lexical verb (LV) and one as the auxiliary verb (AV). The LV in such configurations should be considered the

2 The data sample used for this study attempted for as comprehensive coverage of the subgroups of Chadic as possible, and the sources include both full grammars and short studies on specific topics in the analysis of individual Chadic languages, totaling over 70 languages. I did not modify the data from the original source (except replacing ’ with ʔ, and sometimes added glossing if none was provided in the original). If merited, I on rare occasion offer a different analysis of the data than was offered in the original source. I attempted to be as representative of the different subgroups of the different branches of Chadic as possible. Already dense with information, some decisions were made to cull the original data set presented at the original BICCL meeting and certain language data thus have been omitted here. Thus, we have excluded Bole and cite only one Hausa form among other decisions taken in data reduction.

DOI 10.15460/auue.2025.98.1.394

semantic head, licensing the argument projection frame of the predicate and lexical content of the event predicated. The AV rather specifies or adds event-related and other functional semantic categories or nuances. To illustrate, the English sentences *she has eaten an apple* and *she was eating an apple* predicate of events of 'eating', not of 'having' or 'being', respectively.

There is a crosslinguistic tendency for AVs to be the syntactic head of AVCs (although some variation is attested). Typically, the AV occupies the same position that a simplex verb would in a sentence without an auxiliary. In OV languages, the AV typically follows the LV while in VO languages the AV rather typically precedes the LV.

The functional semantic typology of auxiliary verb constructions in Chadic languages suggests that tense, aspect and mood categories dominate AVCs in these languages. Other categories are only infrequently attested. The most significant variable in the structure of AVCs is the pattern of indexation of inflectional categories on the AV and/or the LV. Often there is just a single pattern attested in a language, but one language can have AVCs of more than one such morphosyntactic configuration too. The best-known pattern cross-linguistically is the AUX-headed structure: inflectional categories are encoded on the auxiliary, and the lexical verb appears in a construction-determined form, often nominalized or Ø-marked, or otherwise is somehow marked as 'dependent'. This in turn generally reflects the syntactic relationship between the two components of the construction (the AV functioning as syntactic head and the LV as syntactic dependent). Formally this dependent status is marked by various means, e.g., a converb form, an infinitive, a verbal noun, a participle, reduplication or prosodically/tonally marked dependency, etc., depending on the specific features of the language as well as the specific terminological conventions of the analytical traditions that predominate in different regions or language families.

From a historical syntactic perspective, AUX-headed AVCs typically arise from verb + clausal complement structures, but can also be serial verb formations in origin too, etc. (see Anderson 2006, 2011 for more details). The headedness characteristics of the LV and AV in AUX-headed AVCs are represented in Table 1.

DOI 10.15460/auue.2025.98.1.394

Table 1. Headedness in AUX-headed AVCs

AUX-headed pattern	**Lexical Verb**	**Auxiliary Verb**
syntactic {'phrasal'}		√ [often]
semantic {'semantic'}	√[3]	
morphosyntactic {'inflectional'}		√

This is the familiar and default structure for AVCs in most well-explored languages. Multiple subtypes of construction-specific dependent forms of LVs (including Ø-marked or bare stem forms) may be found in a given language in various AUX-headed AVCs. For example, *be* in English participates in two different AUX-headed AVCs: with the LV in the *-ing* form it creates a progressive (or imperfect in the past) *(I **am** eat**ing**, I **was** eat**ing**)*, with the LV in the *-ed/-en* form the construction rather has a passive reading *(I **was** eat**en**)*. With this same dependent form of the lexical verb *(-ed/-en)*, when combined with the auxiliary verb 'have', the resulting functional semantics encodes 'perfect' *(I **have** eat**en**)*.

Lexical verbs in AUX-headed AVCs in Chadic languages may be in a nominalized form as in the Bidiya (East Chadic B.1) future-I and future-II forms (ex. 2) or in the East Dangaleat (East Chadic B.1) ventive future (ex. 3).[4]

(2) Bidiya (Alio 1986: 334)

ʔi-kúŋ	*kaatèŋ*	*tí-kùŋ*	*ʔoodyà*
FUT.I-2PL	go:VN	FUT.II-2PL	sleep:VN

'you should go and sleep'

3 In tables 1–4 the notation √ means that this element functions as the head in this structural domain. The notation $\surd_i/\surd_j$ means both features are indexed in this component of the construction while $\surd_i/*_j$ and $\surd_j/*_i$ mean that feature -i and -j show a split distribution (* means the index is lacking on either the LV or AV). In split systems these are mutually exclusive but in split/doubled systems, some categories are mutually exclusively indexed on the LV or the AV in the construction while other categories rather are indexed on both at the same time.

4 Tonal systems in Chadic languages are highly complex, and tone is clearly relevant here. However, I have cited forms with the tonal marking as in the original, although in some cases this may not have been ideal. A more thorough and careful comparison of the tonological properties of the STAMP morph formations and AVCs in Chadic in general needs to be done before a comprehensive and systematic treatment of these phenomena can be achieved.

DOI 10.15460/auue.2025.98.1.394

(3) East Dangaleat (Shay 1999: 234)
ní-k-ke *gàse* *daan* *di*
1PL.EXCL-VEN-2F.OBJ find.VN now only
'we will come and find you'

Such a pattern is also characteristic of the intransitive future in Ga'anda (Central Chadic A.4)[5] as in (ex. 4) and in Za(a)r, also called Sayanci (West Chadic B.3) and Hausa in the progressive or continuous formation (ex. 5 and 6).

(4) Ga'anda (Newman 1971: 20–21)
na-amən *nə́xà-ta* *əssə*
FUT-1PL cook-NMLZ tomorrow
'we will cook tomorrow'

(5) Za(a)r (Sayanci) (Schneeberg 1971: 95)
m-yìgá *nál-gə́nì*
1-IPFV/PROG build-VN
'I am building'

(6) Hausa
ta-na *tafi-ya*
3F-PROG go-VN
'she is going'

Other AUX-headed AVCs in Chadic require rather the lexical verb to remain in a bare stem form. Examples include the progressive in Mpade (Central Chadic C.5) and in Pero (West Chadic A.2) (ex. 7 and 8).

5 Note that future transitive forms show a split pattern with a semi-finite lexical verb. Also, in Ga'anda, the continuous/progressive is realized synchronically as a zero morph (ex. i) but nevertheless requires a nominalized lexical verb and a full/independent/contrastive form of the pronoun.

(i) Ga'anda (Newman 1971: 20–21)
Ø *ŋgət* *nə́xà-ta*
CONT I cook-NMLZ
'I am cooking'

DOI 10.15460/auue.2025.98.1.394

(7) Mpade (Mahamat 2005: 91)
ndá-nè *tò*
PROG-1PL.EXCL return
'we are returning'

(8) Pero (Frajzyngier 1989: 104)
nì-íkka *có* *mín(a)*
1-PROG drink beer
'I am drinking beer'

LEX-headed patterns show the inverse of the AUX-headed pattern: inflectional categories (other than the functional category encoded by the auxiliary element itself) are found on the lexical verb, while the otherwise unmarked auxiliary often remains the syntactic head (Table 2).

Table 2. Headedness in LEX-headed AVCs

LEX-headed Pattern	**Lexical verb**	**Auxiliary verb**
syntactic {'phrasal'}		√ [often]
semantic {'semantic'}	√	
morphosyntactic {'inflectional'}	√	

This is cross-linguistically a relatively uncommon pattern where the auxiliary synchronically behaves like an uninflected element but was historically a verb, grammaticalized as an auxiliary in an AVC. The LEX-headed pattern with an unmarked auxiliary and inflected lexical verb characterizes various non-cognate future AVC forms in a variety of Chadic languages primarily in the Masa and Central Chadic branches. This includes the future in the Masa Chadic language Masa (ex. 9), and the future in the Central Chadic languages Musgu and Hdi (ex. 10 and 11).

(9) Masa (Melis 1999: 93)
ʔàn *mà* *vùl-àŋ-káʔ* *síwì*
1SBJ FUT give-2IND.OBJ-3F.OBJ day.after.tomorrow
'I will give it (FEM) to you the day after tomorrow'

DOI 10.15460/auue.2025.98.1.394

(10) Musgu (Meyer-Bahlburg 1972: 118)

ágá	*mú = gàzì*	*ágá*	*kí = gìzì*
FUT	1SG-come	FUT	2PL = come.PL
'I will come'		'you (PL.) will come'	

àgà	*mù = lúm-kù*
FUT	1SG = eat = 2OBJ

'I will eat you'

(11) Hdi (Frajzyngier & Shay 2002: 197)

dzà'á	*gùy-éy-mú*	*tá*	*vghá*	*màxtsím*
FUT	meet-POT:OBJ-1PL	OBJ	body	tomorrow

'will we meet tomorrow?'

The progressive in Gidar can optionally show a LEX-headed pattern where it varies with an AVC showing a split inflectional pattern (see ex. 20 below). In this Gidar constructional variant, both subject and object are encoded on the lexical verb, i.e., in a configuration of the shape AV LV-OBJ-SUBJ (ex. 12).

(12) Gidar (Frajzyngier 2008: 247)

tà	*wlə-má-nì*
PROG	see-1PL-3PL

'they see us'

In Ga'anda, the perfect, the subjunctive and habitual are all encoded by LEX-headed AVCs. As can be seen in the second form in (ex. 13), more than one auxiliary requiring this configuration can be used simultaneously.

(13) Ga'anda (Newman 1971: 40)

ə	*fəd-úcà-i*	*łəmbərɗa*	*sə*
PRF	beat-2OBJ-1SUBJ	drum	for

'I beat the drum for you'

Ga'anda (Newman 1971: 15)

kə	*łə*	*raka-ən*	*ə*	*walwurca*
SBJV	HAB	run-2	PREP	morning

'you should run in the mornings'

DOI 10.15460/auue.2025.98.1.394

While the LEX-headed pattern is not overly common in general, it nevertheless is attested in various other African languages, e.g., in Central Sudanic Mödö (ex. 14).

(14) Mödö (Central Sudanic) (Persson & Persson 1991: 19)
tí *mó-kɔ̀nyì* *yí*
FUT 1-rescue you
'I will rescue you'

Like any other AVC, the LV in a LEX-headed construction may be in a syntactic dependent relationship with the AV head and this can be indexed by non-finite morphology as in the AUX-headed pattern, e.g., encoded by an infinitive as in Gula Sara (ex. 15).

(15) Gula Sara (Central Sudanic) (Nougayrol 1999: 137)
nɖə́ *kūsá* *gē* *ngá*
PROG INF:eat PL thing
'they/you all are eating'

Co-headed AVCs often arise from core serialized structures or asyndetic coordination but are now mono-clausal AVCs. The distribution of the headedness characteristics of co-headed auxiliary verb constructions is seen in Table 3.

Table 3. Headedness in CO-headed AVCs

Co-headed pattern	**Lexical verb**	**Auxiliary verb**
syntactic {'phrasal'}		√ [often]
semantic {'semantic'}	√	
morphosyntactic {'inflectional'}	√	√

The two most common patterns of the co-headed subtype entail doubling of a TAM marker or of an agreement marker, or both, or indeed doubling of all relevant, obligatory inflectional categories.

Co-headed AVCs are not particularly common in Chadic languages as a rule and the attested formations are largely confined to the Central branch. For example, Central Chadic Mada makes use of various doubly inflected forms, such as the double subject + modal marking in the following co-headed AVC. Note that the effects of the application of harmonic prosodies (and other processes) can make the functionally doubled inflections appear in rather different surface forms (ex. 16).

DOI 10.15460/auue.2025.98.1.394

(16) Mada (Ernst-Kurdi 2016: 85)

nù-uró	*nè-é-dè*	*daf*
1IRR-POT$_{AUX}$	1IRR-POT-cook	millet.coucous

'I am going to cook millet couscous'

Mada (Ernst-Kurdi 2016: 80)

è-mè-né	*à-màá-mbáɗá*	*zzlam*
3IRR-POT-POT$_{AUX}$	3IRR-IRR.POT-read	things

'he might be studying'

Another Central Chadic language showing a doubled subject encoding construction is Muyang (ex. 17).

(17) Muyang (Smith 2010: 103)

á-r(ā)	*á-zòm*	*ɮām*
3-IMM.PRSP$_{AUX}$	3-eat	thing

'he's about to eat something'

As mentioned above, CO-headed AVCs are not very common in Chadic overall, but doubled subject marking is found in core serialized structures in various Chadic languages. It is quite likely that these forms represent a common input source construction for co-headed AVCs in Chadic.

Note that there are, however, certain relatively common formations that might be confused with doubled inflectional patterns: the so-called "intransitive copy pronoun" (ICP) forms in Chadic (Frajzyngier 1977). If bound, ICPs are often in a suffixal series, so are not indexed with prefixed subject pronominals but rather what look like (indirect) object forms or possessive forms, e.g., in Mina (Frajzyngier & Johnston 2011). ICP forms can be found in many (particularly Central and West) Chadic languages, e.g., in West Chadic [A.2] Kushi (ex. 18) or in Central Chadic Gidar (ex. 19). Note importantly that ICPs do not occur with transitive forms as was seen in the 'real' co-headed AVC such as in the Muyang form cited in (ex. 17) above.

(18) Kushi (Batic 2019)

shìnuì	*tà*	*ʔìllɔ̀-jʊ̀*
3PL	FUT	stand.up-3PL.ICP

'they will stand up'

DOI 10.15460/auue.2025.98.1.394

(19) Gidar (Frajzyngier 2008: 143, 141)

sə́ *jáaɓè* *nə̀-dà* *zá-wà*
from Djabe 1-IMM.PST$_{AUX}$ come-1
'I just came from Djabe'

á-nnə́ *sá-w* *á* *já*
FUT-1 be-1 PREP Djabe
'I will be in Djabe'

To be sure the phraseology "intransitive copy pronouns" (ICP) is nearly standard in Chadic linguistics, but it is clearly not an ideal term. These ICP elements do indeed occur in various forms that have other indices of subject, but they also occur in imperatives that lack subject marking in Buwal (Viljoen 2013: 391) where they are clearly not a copy of anything. Note also however that ICP forms in certain South Bauchi languages like Buu/Zaranda and Dott/Zodi (Caron 2006) are not identical with object forms nor possessive pronouns as they are in many other Chadic languages. In West Chadic languages of the Bole-Tangale subgroup, ICP formations seem to occur in languages that have undergone gender levelling, e.g., Kanakuru, Kupto, Kushi, Maha, Pero, Piya, Tangale or Widala (Baldi & Leger 2011: 29). This however does not hold for all Chadic languages.

Split patterns of course do not have a sole inflectional head, but otherwise show the syntactic and semantic headedness characteristics of AUX-headed constructions. This is represented in Table 4.

Table 4. Headedness in SPLIT-headed AVCs

Split pattern	**Lexical verb**	**Auxiliary verb**
syntactic {'phrasal'}		√ [often]
semantic {'semantic'}	√	
morphosyntactic {'inflectional'}	$√_i/*_j$	$*_i/√_j$

There are two common split patterns attested in AVCs, both involving subject indexing split from other indices, specifically a negative marker or an object index. Other patterns are attested but these two sub-types of split patterns constitute the majority of split inflectional systems encountered in AVCs.

A small number of Central Chadic languages show a split in the locus of inflectional indexing of subject vs. object. Thus, in certain AVCs in Gidar, some inflectional sets show the split pattern of subject/

DOI 10.15460/auue.2025.98.1.394

object inflection, with subject marked on the auxiliary and object marked on the lexical verb. This can be seen in variants of the progressive (first form in ex. 20) and the future (second form in ex. 20) in Gidar.

(20) Gidar (Frajzyngier 2008: 154, 263)

mà	*wín*	*tà-t*	*ázgə́l-nì*	*wá-nə̀*	*mpə̀r-kó*
mother	child	PROG-F	feed-3M	FUT-1	chew-2
‘the mother is feeding the baby’				‘I will eat you’	

The Central Chadic language Ga’anda also uses split inflectional patterns in specific AVCs. In the future in Ga’anda with transitive verbs, object is encoded on the lexical verb and subject on the auxiliary (ex. 21).

(21) Ga’anda (Newman 1971: 78, 41)

na-ʔən *raka-ándá-ta*
FUT-1PL.EXCL run-3PL.OBJ-NMLZ
‘we will make them run’

na-i *kwas-ú-ta*
FUT-1 free-2OBJ-NMLZ
‘I will free you’

Note that the LV is nominalized in this structure showing its syntactically dependent status as well. In other words, morphosyntactically, inflectionally, and functionally there is a split, but nevertheless the auxiliary remains in a head-dependent relation with the LV syntactically, marked by the nominalization of the LV.

Causative forms in Gidar show a somewhat unusual split pattern. The causer is encoded as the subject on the lexical verb while the causee appears as the object together with the causative AV element, followed by TAM marking in an atypical configuration of SUBJ-LV AV-OBJ-TAM (ex. 22). Possibly this configuration is evidence that the LV and the AV have become one morphophonological unit in the Gidar causative construction.

(22) Gidar (Frajzyngier 2008: 138)

à-nzá *gà-wə́-kà*
3M-run CAUS-1-PRF
‘he made me run’

DOI 10.15460/auue.2025.98.1.394

Lastly, there are constructions that have similar splits in inflectional headedness as do split AVCs but also simultaneously show doubling of other inflectional categories. These split/doubled formations are somewhat more common in African languages (especially Bantu) than in other macro-regions but are attested globally. The complex nature of the inflectional headedness of split/doubled AVCs is seen in Table 5.

Table 5: Headedness in SPLIT/CO-headed AVCs

Split-Doubled Pattern	**Lexical Verb**	**Auxiliary Verb**
syntactic {'phrasal'}		√ [often]
semantic {'semantic'}	√	
morphosyntactic {'inflectional'}	$\surd_i/*_j$, $*_i/\surd_j$; $\surd_i/\surd_j$	$\surd_i/\surd_j$; $\surd_i/*_j$, $*_{i/\surd j}$

Split/doubled patterns essentially show the same types of patterns as the split pattern but with one category rather doubled, e.g., doubled subject, but another category indexed only once, with for example, single object indexing (on the lexical verb alone in the most commonly attested split/doubled pattern).

Split/doubled patterns are not common in Chadic, but they are attested. The progressive with feminine objects may encode the feminine argument twice on both the lexical verb and the auxiliary in Gidar (ex. 23).

(23) Gidar (Frajzyngier 2008: 160)
ìn-tà-t(=)úlə-tà
1-PROG-3F(=)see-3F
'I see her'

Note that these progressive forms are quasi-univerbated in Gidar, so perhaps these should rather be considered as complex verb forms deriving from an AVC of the split/doubled inflectional pattern (univerbated forms of this sort are discussed in section 4 below). The distinction is only relevant to any given synchronic state of analysis of a specific form, since, as mentioned above, we take a panchronic approach in this study here and thus, even if synchronically a single verbal word in modern Gidar (or on its way to becoming such), the source input construction had to be an AVC of the split/doubled type.

DOI 10.15460/auue.2025.98.1.394

3 Very briefly on the typology of the notion of the verbal 'word' in Chadic languages

Chadic is the largest subgroup in terms of the number of different languages within Afroasiatic, which likely total over 200. While some scholars debate whether three or four branches of Chadic should be reckoned, we adopt here the position of the four-branch hypothesis: East Chadic, Masa Chadic, Central (aka Biu-Mandara) Chadic, and West Chadic. Both the West and East Chadic branches have a handful of subgroups, while the Central Chadic branch has numerous subgroups, although the specific details on which languages belong in which subgroup within Central Chadic remain contested in certain instances.

Syntactically Chadic languages are either verb medial or verb initial[6] – the latter mainly confined to a small number of Central Chadic languages. Since both verb medial and verb initial languages generally have auxiliary verbs that precede lexical verbs, Chadic languages unsurprisingly mainly or typically show AUX V order. Some Chadic languages exhibit TAM-based splits in basic linear syntactic order: Central Chadic Hwana (Hona) has SVO in the progressive and future, but VSO in the past (Frajzyngier 1983: 126) while in Central Chadic Gude, the completive, continuous and potential forms show the order TAM + verb + subject but 'aorist' or 'neutral' aspect forms have the order subject + TAM + verb (Hoskison 1983: 106). Note also in this regard that in Hausa STAMP morphs, the future has the structure AUX = SBJ (e.g., *zân/za nì* [FUT:1SG] *zâ ka* [FUT 2SGM] Newman 2000: 584) as do certain negative forms like the negative continuous, but the vast majority are SBJ + AUX e.g. *ka-na* [2SGM-CONT] (Newman 2000: 575).

Word structure is highly varied across the Chadic family. Certain Central Chadic languages (at least in the analysis of some scholars) have undergone prosodic integration or univerbation of formerly distinct words in a complex verb phrase, that has synchronically resulted in large verbal word complexes with multiple argument indices, verbal specifiers ('extensions'), etc., as in Zulgo (ex. 24) or Ouldeme (ex. 25).

6 VSO languages in Chadic are all in the Central branch, but Frajzyngier (1983: 126) among others considers this to be the original word order pattern in the family.

DOI 10.15460/auue.2025.98.1.394

(24) Zulgo (de Colombel 2003: 44)
avelémèrékérìŋà
a-velé-mèré-ké-rì-ŋà
3.MOD-give.ATTR-1PL.I.O-2OBJ-DIR-3.OBJ
'he will give us something for you'

(25) Ouldeme (de Colombel 1988: 107)
kaŋgàláməsəgètèmekenegè
k-a-ŋgàlá-mə-sagàtày-me-aka-na-egè
2-MOD-AUX:FUT?-again-tire-1PL.OBJ-TLOC-3OBJ-ACHV
'you will again tire yourself out on our behalf because of this'

On the other hand, various forms in specific individual languages present a nearly isolating profile with frequent use of grammatical tone, such as in the following examples from East Chadic Lele (ex. 26), Masa Chadic Musey (ex. 27), Central Chadic Mbara (ex. 28) and West Chadic Mushere (ex. 29):

(26) Lele (Frajzyngier 2002: 173)
gol gé dú
see 3PL 3F
'they saw her'

(27) Musey (Duncanson 1972: 86)
nàm pálà
he DIG:NPST/IPFV
'he digs'

nàm pàlá
he DIG:PST/PFV
'he dug'

(28) Mbara (Barreteau 1987: 47)
mí ná zùm ɬúk
1SG.IPFV FUT eat meat
'I will eat (the) meat'

DOI 10.15460/auue.2025.98.1.394

(29) Mushere (Jungraithmayr & Diyakal 2008 cit. in Schuh 2017: 165)[7]

ka	*náa*	*ká*	*naá*
2SG(M).I	see:PFV	2SG(M).II	SEE:IPFV
'you have seen'		'you will see'	

However, one must approach such data with some caution as criteria for wordhood has not systematically been established and rigorously applied across the Chadic languages, e.g., whether there are distinctions between phonological words and grammatical words, and what those might be distinguished by, i.e., what varying degrees of synthesis are attested and which phono-prosodic domains and phenomena might be used to define which units or indeed whether any units so defined constitute words. Also, one must consider two meta-linguistic factors in determining 'words' in Chadic languages: i) some French-speaking scholars conceive of pre-verbal markers as unbound particles, not prefixes, based on French orthography and likely in part related to this, ii) orthographic conventions adopted for some Chadic languages simply stipulate that the subject markers be written apart, even while phonology tells us they are bound. For example, while subject markers may be written apart in orthographic considerations (and might be called (pro)clitics in specific analyses), nevertheless the harmonic alternations or prosodies and tonal properties of these forms suggest that they must form single phonological words with a following verb, as in Moloko (ex. 30), and thus these subject pronouns seem to really be subject prefixes, and therefore part of the verbal word.

(30) Moloko (Friesen 2017: 208)

7 These should be more properly considered STAMP morph constructions, see section 5 below. Whether such formations should be synchronically considered one word or two simply determines if these are active STAMP morph constructions or have become one word with the former STAMP morph now functioning as a prefixal conjugation marker. This process has happened numerous times in different Chadic languages. Whether a subject pronoun should be considered part of the same word as the verbal element in descriptions of Chadic languages and how to compare these formations across the languages is a major outstanding issue to understanding the diachronic picture. Whether a subject marker is bound or not can indeed vary across different person/number combinations such as the split between 1sg and 3sg in Central Chadic Merey, see (ex. 65) below.

DOI 10.15460/auue.2025.98.1.394

nɛ-tʃɪk-ɛ	*nʊ-tsʊkw-ɔm*
1SBJ-stand-SFX	1SBJ-stand-1PL
'I stand'	'we stand'

These Moloko forms show alternations characteristic of various Chadic languages that reflect the application of palatal(ization) and labial(ization) prosodies (Pearce & Lovestrand 2024), here with leftward spread of the features [front] and [round] respectively, to stems and subject markers from suffixal elements, impacting both consonants and vowels. This suggests that phono-prosodically such forms constitute a single word. To be sure, across the Chadic family, from a morphotactic and tonological perspective,[8] some subject pronominals can be argued to be independent words, some appear to be proclitic, some appear to be fully phono-prosodically integrated prefixes but there are no independent criteria for determining this *per se* applicable to all languages such that like can be compared with like. Also, object indices (direct and indirect) almost invariably follow the verb stem in Chadic languages and are variably described morphotactically as either suffixes or enclitics, again with no independent variable that could be applied across the languages of the family. All these factors must be considered in the discussion below. However, with respect to AVCs, whether a multiverb predicate sequence constitutes one or two 'words' (however defined) at any given point in the history of an individual language simply represents two possible stages in the phonological component of the grammaticalization processes that define AVCs.[9] In summary, much work remains to be done in determining the nature of phono/prosodic and grammatical words and relevant tonology, both within the synchronic description of many individual Chadic languages and in their diachronic analysis as well.

8 Tonal alternations in verb conjugation are complex in Chadic languages and without question they are not irrelevant to this discussion, same for internal vowel changes in stems (ablaut, apophony), but neither of these are discussed *per se* here, just mentioned in passing where necessary, as they each merit their own specialized studies.

9 As noted above, the cline of LV to AV often further entails phono-prosodic integration of the two into a single word that often entails a morphotactic cline of freestanding AV > clitic > affix.

DOI 10.15460/auue.2025.98.1.394

4 Complex verb word forms arising from AVCs in Chadic

Grammaticalization is often accompanied over time by phonological integration of the functional and lexical components of an auxiliary verb construction. Such processes of univerbation yield morphologically complex words, including sometimes even frozen dependent morphology on the lexical verb, whether from an original AUX V (ex. 31) or V AUX (ex. 32) historical configuration.[10]

(31) Swahili (Heine & Reh 1984: 102)
mtoto *a-me-kuja*
1.child 3-PRF-come:FV
'the child has come'

< **mtoto* *a-meele* *ku-ja*
1.child 3-finish:PRF INF-come:FV

(32) Xakas (Turkic; Siberia) (Field Notes)
ojna-p-tɨr-zɨŋ
play-CV-EVID.PST-2SG
'you apparently played'

< **ojna-p* *tur(u)zɨŋ*
play-CV AUX.EVID.PST-2SG

The process of the fusing of former auxiliary verb constructions of such a type into complex verb words is attested in individual Central Chadic languages. Lamang has a wide range of forms that derive from the fusing of original LEX-headed AVCs (ex. 33). Two different modal forms as well as the iterative/durative and the progressive are marked by fused TAM encoding subject pronominals in Lamang.[11] The source of the iterative-durative in Lamang–still in a LEX-headed AVC–was seen above in the Gidar progressive (ex. 20).

10 Indeed the resulting fused structure may preserve an older syntactic configuration in the language as in future forms in Romance languages like French where a form like *chanter-ai* [sing:INF-1SG.FUT] < Vulgar Latin **cantare habeo* 'I will sing' [sing:INF have.1SG.PRS] arose from a V AUX structure while in modern French synchronic AVCs follow the pattern AUX V.

11 Since specific glosses for these forms were not offered in the original source, they will remain unglossed here.

DOI 10.15460/auue.2025.98.1.394

(33) Lamang (Wolff 1983: 143, 145, 170, 164)

	SBJV-II	ITER.DUR	PROG	SBJV-III
1	*a-ts-í*	*tó-ts-í*	*ŋ́-kwòr-ì*	*ká-tsá-yó*
2	*á-tsá-ká*	*tó-ts-ká*	*ŋ́-dzátá-ká*	*ká-tsá-ká*

The Logone subjunctive also likely arose from the fusing of an initial uninflected modal auxiliary (i.e., deriving from an original LEX-headed AVC) with an aorist-marked form of the lexical verb (ex. 34). The source for this univerbation can be seen in the lex-headed configuration in Ga'anda (ex. 21) above.

(34) Logone (Lukas 1936: 47)

ka-də-lo	*ka-i-ló*
SBJV-3F.AOR-come	SBJV-3PL.AOR-come
'she may come'	'they should come'
ka-u-ló	*ka-gə-ló*
SBJV-1AOR-come	SBJV-2AOR-come
'I may come'	'you should come'

5 STAMP morph constructions in Chadic

To be sure among the most noteworthy features of Chadic auxiliary verb constructions is the presence of fused auxiliary and subject markers, sometimes called tensed pronouns, "l'indice de personne aspecto-modal" (Caron 1980: 15), the tense-person complex (Creissels 2005) or STAMP morphs (Anderson 2012, 2015, 2017). STAMP morphs simultaneously encode some canonically verbal functional categories like tense, aspect, mood and/or polarity in addition to the person/number/gender features of (typically subject) participants. They appear to result diachronically from the fusing of subject pronouns and now eroded AVs, sometimes synchronically realized only as tonal alternation (or indeed nothing) and thus lacking any of their original segmental features, while in other cases there is rather more clear evidence of their former auxiliary status. Other than their unusual forms and person/number/gender meanings in addition to typical TAM and polarity functions, STAMP morphs largely function like typical AVs-cum-subject indices and appear in a range of different structural or morphosyntactic configurations. As former AVCs,

DOI 10.15460/auue.2025.98.1.394

it should not be overly surprising that a given specific STAMP morph construction may require LVs to appear in a construction-specific form, that can at times require non-finite (or semi-finite) forms, like any AVC might (have done). STAMP morphs are thus a possible intermediate stage in a grammaticalization continuum of AVCs where the form functioning as a subject pronoun and a TAM marker was historically reinterpreted as a single word (resulting from the fusing of an original subject + AV formation).[12]

STAMP morphs occur in all four subgroups of Chadic, and across several subgroups within these. They are particularly common in certain West Chadic languages. STAMP morphs in Chadic languages constitute a sizable variety of forms when viewed as a group. Whether these formations are old or can be attributed to any earlier stages in the history of Chadic is the subject of our present research as even closely related languages can show seemingly non-cognate systems.

The simplest systems are probably found in languages like Miya of the West Chadic B.2 branch (ex. 35), in Mpade and Mbara of the Central Chadic branch (ex. 36, ex. 37) or in Migama and Gali/Miltu of the East Chadic branch (ex. 38, ex. 39). Here STAMP morphs occur within an aspectual system contrasting perfective with imperfective (Mbara, Miya, Mpade) or in a modal system contrasting conditional vs. indicative (Migama) or in a tense(-cum-modal) system contrasting future and non-future in Gali/Miltu. In all these languages, the STAMP morphs take unmarked lexical verb stem forms.

(35) Miya (Schuh 2001: 444)

tə́	*má*	*mara*	*zhaak-uw*
3.IPFV	NEG	donkey	get-CONEG

'he will not get a donkey'

12 Thus, STAMP morph constructions can be in any structural configuration that an auxiliary verb construction can be in. So they may be AUX-headed, LEX-headed, split, co-headed or show split-doubled distribution in their inflectional patterns, just as AVCS can. STAMP morphs are therefore just a possible intermediate stage in the cycle of freestanding full words to bound/univerbated structures that grammaticalized AVCs might pass through (i.e., formations that will fuse together to become complex words). The STAMP morph stage is thus one in which a subject index (whether bound or free) and a functional auxiliary element have univerbated into one phonoprosodic complex. This in turn has entered into a construction with at least a lexical verb , which may also index from zero up to many other potential grammatical indices not encoded by the STAMP morph.

DOI 10.15460/auue.2025.98.1.394

à *már* *mà* *zhaak-úw*
3.PFV donkey NEG get-CONEG
'he didn't get a donkey'

(36) Mpade (Mahamat 2005: 88; 53)
liwdò *wò* *lù* *hò* *dó* *kídá* *rò*
yesterday 1.PRF come house DEM work DEM

gídí *dó* *nò* *gè*
for.her DEM 3F.PRF finish
'when I returned yesterday, her work was already finished'

sí *làkè* *ù* *dì* *kàsúgú* *gó* *ábà* *ŋgìnè*
day every 1SG.IPFV go walking with father 1POSS
'I go walking every day with my father'

(37) Mbara (Tourneaux 1978: 30)

á *sér*	*à* *sér*
3M.IPFV crash	3M.PRF crash
'he crashes, is crashing'	'he has crashed'

(38) Migama (Jungraithmayr & Abakar 1992: 37)

ná *ʔàsáa*	*náa* *ʔàsáa*
1.IND come	1.COND come
'I come, am coming'	'if I come'

(39) Gali/Miltu (Jungraithmayr & Peust 2019: 224)

nə́ *tíyé* *màn*	*nə́ə̀* *tē/tə̄* *màn*
1.NFUT eat.PFV food	1.FUT eat.FUT food
'I ate/have eaten'	'I will eat'

The future STAMP morph series in East Dangaleat (East Chadic) requires a verbal noun form of a lexical verb it occurs with (ex. 40). Highly unusual for Chadic is the fact that the STAMP morph also serves as the host of object agreement morphology (ex. 41) if the specific verbal configuration merits such (i.e., a transitive verb with overt object indexing).

DOI 10.15460/auue.2025.98.1.394

(40) East Dangaleat (Shay 1999: 212)
kíi *tálé*
2M.FUT see:VN
'you will see'

(41) East Dangaleat (Shay 1999: 213)
maa *di* *an-dyìŋ* *dɛŋ̀*
why only FUT.1SG-2M.OBJ kill/finish.VN
'why should I kill you?'

In West Chadic A.3 Angas, STAMP morphs may occur with various other auxiliary elements as well in complex multipart AVCs. Compare the two phrases in (ex. 42).

(42) Angas[13] (Burquest 1973: 38)
ŋâː *jì* *ŋá* *mɛ́t* *jì*
1.COMPL come 1.NPRS FUT come
'I have come' 'I will come'

Many Chadic languages show more developed systems of STAMP morphs. East Chadic Kwang (ex. 43), West Chadic (A.3) Mushere (ex. 44) and Montol (ex. 45) and Central Chadic Kirya-Konzel (ex. 46) each have at least three distinct series.

In Kwang, the perfective series also requires perfective stem-ablaut and suffixation on the lexical verb, compare (ex. 43) with (ex. 44). This type of patterning is not uncommon in East Chadic languages, and it may also be found in West Chadic languages as well.

(43) Kwang (Ebert 1987: 65)
ú *sé* *wa* *sé*
3M.AOR drink 3.FUT/HAB drink
'he drank' 'he has drunk'

13 When glossing was lacking in the original source, I have tried to add this if possible. Naturally some errors may have crept in during such and these are my responsibility, and I hope nothing terribly inaccurate has been proposed.

DOI 10.15460/auue.2025.98.1.394

(44) Kwang (Ebert 1987: 66)

ù	*sàá-n*
3M.PRF	drink.PRF-PRF

'he will drink'

The three series of Mushere encode PROG, PST, FUT according to the analysis of Jungraithmayr & Diyakal (2013); see (ex. 45).[14]

(45) Mushere (Jungraithmayr & Diyakal 2013: 304)

	PROG	PST	FUT
1	*an*	*a*	*á*
2M	*nga*	*ka*	*ká*
2F	*nji*	*yi*	*yí*
3M	*ni*	*ní*	*dǐ*
3F	*ni*	*ní*	*dǐ*

A nearly identical functional opposition is found West Chadic Montol, here contrasting a PRF, a FUT and a PROG series of STAMP morphs (ex. 46); offered here are the singular forms.[15]

(46) Montol (Jungraithmayr 1964: 169–70)

	PRF	FUT	PROG
1	*a*	*á laa*	*hăndî*
2M	*gə*	*gə́*	*gaa*
2F	*zhi*	*zhí*	*zhîi*
3	*ni*	*ká*	*nîi*

Three sets of STAMP morphs are also attested in Central Chadic Kirya-Konzel. Similar to Mushere, Kirya-Konzel contrasts IPFV, PFV, and FUT series of STAMP morphs (ex. 47).

14 As mentioned above, it is always possible we are dealing with bound prefixes rather than freestanding STAMP morphs in any Chadic language presented here, if analysts chose to write subject encoding elements as separate words not as prefixes. But these two options are just two stages in the phono-prosodic history of such formations. As mentioned previously orthographic considerations do not always reflect phonological reality.

15 To be sure, while not yet determined, it is entirely plausible that the PRF series of Montol and PST series of Mushere are cognate.

DOI 10.15460/auue.2025.98.1.394

(47) Kirya-Konzel (Blench & Ndamsai 2009: 82)

ǯò	*hwyi*	*câwcâw*	*wàyí*	*kì*	*kə́rí*
1.IPFV	run	quickly	1.FUT	beat	dog
‘I am running quickly’			‘I will beat the dog’		

ǯè	*kì*	*kə́rí*
1.PFV	beat	dog

‘I beat the dog’

The Ron Chadic language Fyer (West Chadic A.4) has many STAMP morph series. The singular forms of the basic seven sets (with the corresponding tonal melody or projected floating tone on the verb stem) are presented in (ex. 48). There are two sets of series, one with short vowels (and also all high in the SG series except 2M) and largely requiring lexical verbs to bear a low tone (the future can project either L or H tone onto the verb), and the other with long low vowels in the STAMP morph and projecting an H tone onto the lexical verb.

(48) Fyer (Jungraithmayr 1970: 44)

	PRF	HAB	PROG	SUB	AOR	SBJV	FUT
1	*yáà*	*yaá*	*yaâ*	*yáa*	*yí*	*yì*	*yi*
2M	*háà*	*haá*	*haâ*	*háa*	*há*	*hà*	*ha*
2F	*sháà*	*shaá*	*shaâ*	*sháa*	*shí*	*shì*	*shi*
3M	*máà*	*maá*	*maâ*	*máa*	*mí/mú*	*mì/mù*	*m(i/u)/n*
3F	*táà*	*taá*	*taâ*	*táa*	*tí*	*tì*	*ti*
+ tone of LV	H	H	H	H	L	L	L/H

Some examples of forms with the Fyer STAMP morphs and their associated tonal projections can be seen in (ex. 49).

(49) Fyer (Jungraithmayr 1970: 45)

yáà	*gán*	*yaá*	*gán*	*yaâ*	*gán*	*yáa gán*
1.PRF	go	1.HAB	go	1.PROG	go	1.TEMP go
‘I have gone’		‘I used to go’		‘I am going’		‘when I went’

DOI 10.15460/auue.2025.98.1.394

yí	*nyì*	*yì*	*nyì*	*yi*	*nyí*
1.AOR	do	1.SBJV	do	1.FUT	do
‘I do/did’		‘I may do’		‘I will do’	

As the above formations instantiate, STAMP morph constructions are often in an AUX-headed configuration, but they may also occur in different structural configurations in Chadic. Thus, in East Chadic Mawa, aspect is encoded in the STAMP morph (signaled tonally) and affixally on the lexical verb as well, in a doubled aspect indexing form (ex. 50).

(50) Mawa (Jungraithmayr 1981: 55)

no	*waany~wáánya-ŋ*	*nó*	*waany-ê*
1.IPFV	RDPL~open-IPFV	1.PRF	open-PRF
‘I open’		‘I opened’	

West Chadic Tangale (West Chadic A.2), like Mawa, also doubly marks aspect in STAMP morph constructions, once in the STAMP morph itself and once via a suffix on the lexical verb, as in the following perfect and continuous forms (ex. 51).[16]

(51) Tangale (Jungraithmayr 1971: 29)

n	*wee-go*	*naŋ*	*sol-i*
1.PRF	see-PRF	1.CONT	pull-CONT
‘I have seen’		‘I am pulling’	

In Batna (Masa Chadic), perfective and imperfective STAMP morphs can be distinguished in a small number of person/number combinations. These STAMP morphs appear with a verb stem in an aspect-specific tonal form, L for imperfective, H for perfective (ex. 52).

(52) Batna (Jungraithmayr 1978: 7)

ndaá	*pum*	*ndaa*	*púm*
3(F).PFV	beat.IPFV	3(F).PFV	beat.PFV
‘she beats’		‘she has beaten’	
mbâa	*pum*	*mbáà*	*púm*
1PL.IPFV	beat. IPFV	1PL.PFV	beat.PFV
‘we beat’		‘we have beaten’	

16 Note that a 1PRF form *n* is found in the grammatical systems of several Niger-Congo languages (Anderson 2012, 2017).

DOI 10.15460/auue.2025.98.1.394

naá	*pum*	*naa*	*púm*
1.IPFV	beat.IPFV	1.PFV	beat.PFV
'I beat'		'I have beaten'	

Negative and positive forms are contrasted in the perfective series in Bachama (Central Chadic A.1 or "Bata" group), with different aspect-cum-polarity suffixes on the lexical verb (ex. 53):

(53) Bachama (Carnochan 1970: 86; 85)

tèé	*zùm-i*	*hye*	*zùm-o*
2.NEG.PRF	eat-NEG.PRF	2.PRF	eat-PRF
'you didn't eat it'		'you ate it'	

Other individual Chadic languages may show varied inflectional patterns involving STAMP morph constructions. For example, the West Chadic B.2 language Pa'a has unmarked verb stems with the STAMP morph in the perfect series (marked by H tone on the STAMP morph) (ex. 54);

(54) Pa'a (Jungraithmayr 1966/67: 198)

má	*sà*	*ʔú*	*sà*	*ná*	*mìy*
1.PRF	drink	2.PRF	drink	3M.PRF	die
'I have drunk'		'you have drunk'		'he has died'	

But the future form in Pa'a is marked by a compound AV. This consists of an L tone marked STAMP morph that combines with an inflected auxiliary verb (< 'go') which marks its own subject by suffixes, followed by a dependent lexical verb form. Thus, perhaps it is something like a doubled auxiliary formation with double-subject encoding (for some person/number combinations at least) in origin, with the second auxiliary requiring a verbal noun form of the LV. This auxiliary is intransitive etymologically and the subject marking is suffixal here. Given how rare true doubled subject marking is in Chadic AVCs and that suffixal subject indexing is relatively marked in West Chadic, perhaps this really should be considered a STAMP morph + Aux-ICP + LV_{VN} formation from an historical or diachronic perspective. Thus, this form does not actually exhibit a doubled subject formation as such. The original meaning of the future auxiliary in Pa'a is 'go' so it meets the conditions necessary to have once been an ICP formation. See Pa'a examples in (ex. 55).

DOI 10.15460/auue.2025.98.1.394

(55) Pa'a (Jungraithmayr 1966/67: 198)

wù	*t-ú*	*mìyà-w*	*mà*	*t-ə́n*	*mìyà-w*
2.FUT	FUT-2	die-VN	1.FUT	FUT-1	die-VN
'you shall die'			'I shall die'		

nà	*t-á*	*mìyà-w*
3M.FUT	FUT-3M	die-VN

'he shall die'

STAMP morph forms can also occur in the split subject/object pattern where the STAMP morph encodes subject and tense, and the lexical verb the object as in West Chadic Polci (ex. 56) or in Central Chadic Mofu-Gudur (ex. 57).

(56) Polci (Caron 2008: 153)

Gărbà	*kən*	*ndʒaŋ*	*slo:*	*wú*	*de*	*kə*
Garba	COP	slit.throat	animal.meat	0ACC	INJ	2:AOR

fŭ:-m
say-1

'If Garba slays a beast for meat, tell me, tell me'

(57) Mofu-Gudur (Pohlig 1992: 4)

fá	*tá-ka*	*ɗáf*
PROG.3	prepare-2.IO	food

'she is preparing you food'

When objects are indexed in transitive verbs in the Pa'a future, a split/doubled pattern is found with the subject encoded in the STAMP morph (by definition) and with the object appearing on a semi-finite dependent/nominalized verb.

DOI 10.15460/auue.2025.98.1.394

(58) Pa'a (Jungraithmayr 1966/67: 198)

mà	*t-án*	*góó-sù*	*mà*	*t-án*	*góó-sə́ŋ*[17]
1.FUT	FUT-1	give-3.OBJ	1.FUT	FUT-1	give-3PL.OBJ
'I shall give him/her'			'I shall give them'		

6 Prefix conjugations arising from STAMP morph constructions in Chadic

Prefix conjugations in Chadic languages can in many cases be demonstrated to result from the fusing of an original STAMP morph formation and a following lexical verb. That is, these forms show two different processes of univerbation: an original one to create the STAMP morph and a secondary process of fusing of the already existing STAMP morph and the former lexical verb in the STAMP morph construction into a conjugational prefix. This phenomenon is found in at least three subgroups of Chadic (possibly not Masa Chadic), but it is especially common in some subgroups of Central Chadic. For example, Mbuko offers a straightforward and simple instantiation of a prefixal conjugation deriving from a fused STAMP morph structure of the AUX-headed type. Here STAMP morphs have been univerbated with the lexical verb, and the perfect is marked by an additional perfect suffix (ex. 59).

(59) Mbuko (Gravina 2001: 7)

nə-zlāmbāl-ák	*nà-zlàmbál*
1PRF-throw-PRF	1ANT-throw:ANT
'I have thrown'	'I threw'

The Central Chadic language Zulgo also contrasts two sets of forms, one perfective and the other imperfective in prefixed conjugational series that derive from fused STAMP morph constructions (ex. 60).

17 It is not clear what Jungraithmayr intends with the diacritic notation on *ə́*. Clearly the ´ part indicates high tone of the vowel, but what he intended to convey by the symbol *ə̂* remains opaque. It is clearly non-tonal as the form is not simultaneously high toned plus falling tone. According to Skinner (1979) the language does have a falling tone and [*ə*] is an allophone of */a/* medially. Moreover, Skinner (1979: 41) represents the form glossed FUT-1 here as *tín* not with a schwa and the 3PL possessive suffix here functioning as an object index is transcribed by her as *-sìn*, with [*i*] an allophone of */i/*.

DOI 10.15460/auue.2025.98.1.394

(60) Zulgo (Haller et al. 1981: 47–8; 31, 49)

í-dé-áha = á *dzékwiŋ*
1.IPFV-go.IPFV-TLOC = GEN.EVT then
'I will go there then'

ì-dá-ára *á* *máp̀rà-ákà-kár-á* *mbákum*
1PFV-go.PFV-CLOC to see-on-2OBJ-2CLOC today
'I came to see you today'

á-víl-iŋ *mendzikwir*
3PFV-give-1.I.OBJ chicken
'he gave me a chicken'

à-dé-áha-áɗəm-íŋá
3IPFV-go.IPFV-TLOC-into.it-GEN.EVT
'he is going into it'

This same pattern appears to lie at the origin of the prefixal agreement series in Central Chadic Buduma (ex. 61).

(61) Buduma (Awagana 2003: 3)

yágə̀nàỳ *yá-fì* *mə̀dáy*
children 3PL.PFV-hit RCP
'the children are hitting each other'

yágə̀nàỳ *má* *yə̀-fábí*
children DET 3PL.IPFV-hit.PL:PSV
'the children engage in hitting without ceasing'

Logone also has two separate series of verb forms historically derived from STAMP morphs. Thus, the aorist series and the preterite series show distinct forms for all person/number combinations.

(62) Logone (Lagwan) (Lukas 1936: 38–39)

ú-l	*u-ká*	*wá-l*	*wā-ka*
1AOR-go	1AOR-say	1PRET-go	1PRET-say
'I go'	'I say'	'I went'	'I said'

Gashua Bade (West Chadic) likewise shows a system with prefixal agreement series for imperfective and perfective both derived from univerbated STAMP morph constructions (ex. 63):

DOI 10.15460/auue.2025.98.1.394

(63) Gashua Bade (Ziegelmeyer 2013: 6)

naa-bd-aali	*nə̂-bd-acî*
1IPFV-ask-GEN3M	1PFV-ask-3M.OBJ
'I will ask him'	'I asked him'

Some Chadic languages show distinct patterns in different inflectional sets. For example, Central Chadic Wandala has fused STAMP morphs deriving from source/input constructions that had distinct inflectional patterns. One is synchronically prefixal (future) and the other synchronically is an infix appearing in the middle of what appears to be a reduplicated stem (preterite) (ex. 64).

(64) Wandala (Lukas 1937: 118–9)

1SG.FUT *yá-*

1SG.PRET *-ném-* + RDPL

	'buy'	'bring'
1SG.FUT	*yá-ʃʃùkwà*	*yá-sánsà*
1SG.PRET	*ʃèkwà-nép-ʃúkwé*	*sà-nén-sà*

Other variation patterns are attested too. Thus, in past forms in the Central Chadic language Merey, the first-person singular appears to be a freestanding STAMP morph structure with the lexical verb in an unmarked form in an AUX-headed pattern, while the corresponding third person masculine singular form is rather a fused structure with an inflectional prefix deriving from a STAMP morph construction. Compare 1sg forms with their corresponding 3SG forms (ex. 65):

(65) Merey (Gravina 2007: 8)

na	*zal*	*a-zal*
1.PST	call	3.PST-call
'I called'		'he called'

na	*ge*	*a-ge*
1.PST	do	3.PST-do
'I did'		'he did'

The East Chadic language Mokilko (Jungraithmayr 1990) shows an extremely complex system. This combines a set of seven basic TAM stem form subtypes (listed as I–VII in ex. 66), with many subclasses of stem shape, e.g., a perfective and two variants of the imperfective

DOI 10.15460/auue.2025.98.1.394

with their own vocalic (and even consonantal) shapes and tonological pattern; see (ex. 66).

(66) Mokilko (Jungraithmayr 1990: 44–45)

	I	VII	II
	AOR	FUT.ANT	PROG
‘eat’	*ʔîm-í*	*ʔîm-áà*	*ʔômɓ-ó*
‘look for’	*díʔ-é*	*díʔ-áà*	*déʔ-ú*
stem class	PFV	PFV	IPFV.II

	III	IV	V	VI
	SBJV	HORT	PAST	FUT
‘eat’	*ʔóòm-í*	*ʔóòm-ó*	*ʔóòm-ò*	*ʔóòm-áà*
‘look for’	*déʔ-é*	*déʔ-ó*	*déʔ-ò*	*déʔ-áà*
stem class	IPFV	IPFV	IPFV	IPFV

Any of these stem shapes that encode the lexical content and (im)perfectivity and other TAM concepts can be inflected in eleven different prefix conjugational series (listed as A–K in ex. 67).

(67) Mokilko

	A	B	C	D
3M	*y(íi)-*	*yà(à)-*	*yáà-*	*yáàd(ì)-*
3P	*ʔán-*	*ʔând(ì)-*	*ʔànáà-*	*ʔànáàd(ì)-*

	E	F	G	H
3M	*yáat(í)-*	*yáàt(í)-*	*yât(í)-*	*yíi-*
3P	*ʔânt(í)-*	*ʔànâtt(í)-*	*ʔanâtt(í)-*	*ʔáŋ-*

	I	J	K
3M	*yíid(í)-*	*yíit(í)-*	*yít(í)-*
3P	*ʔánd(í)-*	*ʔánt(í)-*	*ʔânt(í)-*

Originally these prefixed inflectional sets were likely STAMP morph constructions, but in the present state of Mokilko these all became fused into a highly synthetic structure (ex. 68).

DOI 10.15460/auue.2025.98.1.394

(68) Mokilko (Jungraithmayr 1990: 36, 43)

yáàd-îm-í	*yà-ʔîm-í*
3M.D-eat.AOR-PRF	3M.B-eat.AOR-PRF
'he ate it again'	'he has finally eaten'

yáàt-óòm-í
3M.E-eat.IPFV-SBJV
'so that he eats'

It is important to note here that more than a single source construction can result in a complex verb form in the synchronic grammar of specific Chadic languages. Specifically, this may combine forms derived from a STAMP morph construction with forms derived from an AVC. For example, the Gulfei Kotoko (Central Chadic) preterite derives from a STAMP morph construction and synchronically functions as a prefixal conjugation (ex. 69), while the future in the same language rather derives from a fused LEX-headed AVC structure, not a STAMP morph. It is therefore likely in this latter case that what synchronically appears to be subject marking indexed by an infix (ex. 70), is etymologically a prefix on the lexical verb from an original LEX-headed AVC.

(69) Gulfei Kotoko (Lukas 1937: 153)

gá-yím	*wà-yím*
2.PRET-eat	1.PRET-eat
'you ate'	'I ate'

(70) Gulfei Kotoko (Lukas 1937: 153)

na-g-yim	*na-u-yim*
FUT-2-eat	FUT-1-eat
'you will eat'	'I will eat'

7 Conclusion

Chadic languages attest different inflectional configurations in AVCs and forms derived historically from such. AUX-headed and LEX-headed inflectional patterns are most common, while split, co-headed and split/doubled patterns are infrequently found. Likely the most

DOI 10.15460/auue.2025.98.1.394

characteristic feature of forms derived historically from AVCs in the Chadic family is the many varied series of subject indices that simultaneously encode TAM and polarity meanings – the so-called STAMP morphs – that generally arose from fusing of subject pronouns and auxiliaries. Both prefix and infix agreement series may arise in Chadic languages from the univerbation of STAMP morphs with lexical verbs into large complexes.[18] The original inflectional pattern of the STAMP morph construction is often reflected in these verbal complexes. Less commonly, other complex verb forms can be demonstrated to have derived also from the univerbation of former AVCs into larger complexes in various Chadic languages. The fact that STAMP morphs have been generated and recruited into prefixal agreement conjugations over and over in the history of different subgroups of Chadic helps explains why these types of constructions in Chadic do not appear to be cognate with prefixal morphology of related Afroasiatic language groups and thus are not reflexes of proto-Afroasiatic prefixes. Rather they reflect tendencies found across the Macro-Sudan Belt for subject markers to coalesce into STAMP morphs with eroded auxiliaries and for these to eventually become prefixal conjugational markers. Similar phenomena have also happened in Niger-Congo and Central Sudanic languages (Anderson 2012, 2015, 2017) in addition to Chadic.

18 This does not include forms in various Central Chadic languages where what appears to be an infixed subject index occurs between a base and reduplicant (or reduplicant and base) in certain reduplicated TAM formations.

DOI 10.15460/auue.2025.98.1.394

Abbreviations

.I	Type-I
.II	Type-II
√	Verb Root
1	1st person
2	2nd person
3	3rd person
ACHV	Achievement
ANT	Anterior
AOR	Aorist
ATTR	Attributive
AUX	Auxiliary
AV	Auxiliary verb
AVC	Auxiliary Verb Construction
CAUS	Causative
CLOC	Cislocative
COMPL	Completive
COND	Conditional
CONEG	Conegative
CONT	Continuous
D.O.	Direct Object
DEM	Demonstrative
DIR	Directional
DIR.OBJ	Direct Object
DUR	Durative
EXCL	Exclusive
F	Feminine
FUT	Future
FV	Final Vowel
GEN	Genitive
GEN.	
EVT	General Event
HAB	Habitual
ICP	Intransitive Copy Pronoun
IMM	Immediate
IN	Inclusive
IND	Indicative
IND.	
OBJ	Indirect Object
INF	Infinitive
I.O	Indirect Object
IPFV	Imperfective
IRR	Irrealis
ITER	Iterative
LEX	Lexical
M	Masculine
MOD	Modal
NEG	Negative
NMLZ	Nominalizer
NPRS	Non-Present
OBJ	Object
PFV	Perfective
PFX	Prefix
PL	Plural
POT	Potential
PREP	Preposition
PRET	Preterite
PRF	Perfect
PROG	Progressive
PRS	Present
PRSP	Prospective
PST	Past
RDPL	Reduplication
SBJV	Subjunctive
SBJ	Subject
SFX	Suffix
SUB	Subordinate
TAM	Tense-Mood-Aspect
TEMP	Temporal(is)
TLOC	Translocative
VEN	Ventive
VN	Verbal Noun

DOI 10.15460/auue.2025.98.1.394

References

Alio, Khalil. 1986. *Essai de description de la langue Bidiya du Guéra (Tchad). Phonologie-Grammaire.* (Marburger Studien zur Afrika- und Asienkunde. Serie A: Afrika Bd. 45). Berlin: Reimer.

Anderson, Gregory D. S. 2006. *Auxiliary verb constructions.* Oxford: Oxford University Press.

Anderson, Gregory D. S. 2011. Auxiliary verb constructions in the languages of Africa. In *Studies in African Linguistics* 39(1/2). 1–342.

Anderson, Gregory D. S. 2012. S/TAM/P morphs in the history of Benue-Congo and Niger-Congo conjugation. Paper presented at Niger-Congress, Paris, September 2012.

Anderson, Gregory D. S. 2015. STAMP morphs in Central Sudanic languages. In Angelika Mietzner & Anne Storch (eds.), *Nilo-Saharan: Models and descriptions.* Köln: Rüdiger Köppe Verlag. 151–167.

Anderson, Gregory D. S. 2017. STAMP morphs in the Macro-Sudan Belt. In Doris Payne, Sara Pacchiarotti & Mokaya Bosire (eds.), *Diversity in African languages.* Berlin: Language Sciences Press. 513–539.

Awagana, Ari. 2003. La Pluralité verbale en Buduma. In H. Ekkehard Wolff (ed.), *Topics in Chadic linguistics. Papers from the 1st Biennial International Colloquium on the Chadic Languages, Leipzig, July 5-8, 2001.* Köln: Rüdiger Köppe Verlag. 1–20.

Baldi, Sergio, & Rudolf Leger. 2011. Some diachronic observations on gender and number in Bole-Tangale languages. In Doris Löhr & Ari Awagana (eds.), *Topics in Chadic linguistics VI: Papers from the 5th Biennial International Colloquium on the Chadic Languages.* Köln: Rüdiger Köppe Verlag. 23–33.

Barreteau, Daniel. 1987. Du mbara au mofu-gudur: approches comparées des systèmes verbaux. In Herrmann Jungraithmayr (ed.), *Langues tchadiques et langues non-tchadiques en contact en Afrique Centrale.* Paris: SELAF. 37–60.

Batic, Gian Claudio. 2019. Verb classes and TAM system in Kushi (Nigeria). In Dymitr Ibriszimow, Henry Tourneux & H. Ekkehard Wolff (eds.), *Topics in Chadic linguistics X*. Köln: Rüdiger Köppe Verlag. 67–84.

Blench, Roger & Antony Ndamsai. 2009. Introduction to Kirya-Konzel: A central Chadic language of Eastern Nigeria. In Eva Rothmaler (ed.), *Topics in Chadic linguistics V: Comparative and descriptive studies. Papers from the 4th Biennal International Colloquium on the Chadic*

DOI 10.15460/auue.2025.98.1.394

Languages. Bayreuth, October 30–31, 2007. Köln: Rüdiger Köppe Verlag. 75–85.

Burquest, Donald A. 1973. 1980 (1977). Angas. In Mary E. Kropp-Dakubu (ed.), *West African language data sheets. Volume 1*. Accra: West African Linguistic Society. 35–43.

Bybee, Joan L., Revere Perkins & William Pagliuca. 1994. *The evolution of grammar. Tense, aspect, and modality in the languages of the world.* Chicago: University of Chicago Press.

Bybee, Joan L. & Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1). 51–103.

Carnochan, Jack. 1970. Categories of the verbal piece in Bachama. *African Language Studies* 11. 81–112.

Caron, Bernard. 1980. Aoristique et subjonctif haoussa. In Herrmann Jungraithmayr (ed.), *Études sur le subjonctif en tchadique.* (Africana Marburgensia Sonderheft 4). Marburg: Lahn. 13–39.

Caron, Bernard. 2006. South-Bauchi West Pronominal and TAM Systems. In Bernard Caron & Peter Zima (eds.), *Sprachbund in the West African Sahel.* Louvain: Peeters. 93–112.

Caron, Bernard. 2008. La structure énonciative des subordonnées conditionnelles. In Bernard Caron (ed.), *Subordination, dépendance et parataxe dans les langues africaines.* Louvain: Peeters. 145–158.

Creissels, Denis. 2005. A typology of subject and object markers in African languages. *In F. K. Erhard Voeltz (ed.), Studies in African linguistic typology. Amsterdam: Benjamins. 43–70.*

De Colombel, Véronique. 1988. Mode, temps, aspect en tchadique. Mode, temps, aspect en tchadique. In Nicole Tersis & Alain Kihm (eds.) *Temps et aspects.* Paris: Peeters/SELAF.103-115.

De Colombel, Véronique. 2003. Grammaticalisation et ordre des unités dans dix langues du groupe tchadique-central. In. H. Ekkehard Wolff (ed.), *Topics in Chadic Linguistics. Papers from the 1st Biennial Colloquium on Chadic Languages (Leipzig, July 5–8, 2001)*. Köln: Rüdiger Köppe Verlag. 43–51.

Duncanson, Robert. 1972. Interpreting tone in the Musey language. Gounou Gaya: Mission Protestante Ms.

Ebert, Karen H. 1987. A first comparison of Kera and Kwang. In Herrmann Jungraithmayr (ed.), *Langues tchadiques et langues non-tchadiques en contact en Afrique Centrale.* Paris: SELAF. 61–70.

Ernst-Kurdi, Eszther. 2016. *Tense, aspect and mood in Mada, a Central Chadic language.* Gloucester: Redcliffe College MA Thesis.

DOI 10.15460/auue.2025.98.1.394

Frajzyngier, Zygmunt. 1977. On the intransitive copy pronouns in Chadic. *Studies in African Linguistics* 7. 73–84.
Frajzyngier, Zygmunt. 1983. Marking syntactic relations in Proto-Chadic. In H. Ekkehard Wolff & Hilke Meyer-Bahlburg (eds.), *Studies in Chadic and Afroasiatic linguistics*. Hamburg: Helmut Buske. 115–138.
Frajzyngier, Zygmunt. 1989. *A grammar of Pero*. Berlin: Dietrich Reimer.
Frajzyngier, Zygmunt. 2002. System interaction in the coding of modality. In Sjef Barbiers, Wim Wurff, & Frits Beukema (eds.) *Modality and its interaction with the verbal system. Amsterdam: Benjamins.* 165–84.
Frajzyngier, Zygmunt. 2008. *A grammar of Gidar.* (Research in African Studies 13). Frankfurt: Peter Lang.
Frajzyngier, Zygmunt & Eric Johnston. 2011. *A grammar of Mina.* Berlin: Mouton de Gruyter.
Frajzyngier, Zygmunt & Erin Shay. 2002. *A grammar of Hdi*. Berlin: Mouton de Gruyter.
Friesen. Dianne. 2017. *A grammar of Moloko*. Language Science Press.
Gravina, Richard. 2001. *The verb phrase in Mbuko*. Yaoundé: SIL.
Haller, Beat, Sylvia Hedinger & Ursula Wiesemann. 1981. The verbal complex in Zulgo. In Ursula Wiesemann (ed.) *Chadic language studies in Northern Cameroon*. (Africana Marburgensia Sonderheft 5). Marburg: Lahn. 17–54.
Heine, Bernd 1993. *Auxiliaries. Cognitive forces and grammaticalization.* New York: Oxford University Press.
Heine, Bernd & Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske Verlag.
Hoskison, James Taylor. 1983. *A grammar and dictionary of the Gude language*. Columbus: Ohio State University Ph.D. dissertation.
Jungraithmayr, Herrmann. 1964. Materialien zur Kenntnis des Chip, Montol, Gerka und Burrum/Südplateau, Nordnigerien. *Afrika und Übersee* 48. 161–182.
Jungraithmayr, Herrmann. 1966/67. Specimens of the Pa'a ("Afa") and Warja languages with notes on the tribes of Ningi Chiefdom (Bauchi Province, Northern Nigeria). *Afrika und Übersee* 50. 194–205.
Jungraithmayr, Herrmann. 1970. *Die Ron-Sprachen: Tschadohamitische Studien in Nordnigerien.* (Afrikanistische Forschungen 3). Glückstadt: J. J. Augustin.
Jungraithmayr, Herrmann. 1978. The Zime dialect cluster ('Kado', 'Dari') in southern Chad: Its verbal aspect system. *Afrika und Übersee* 61. 1–27.

DOI 10.15460/auue.2025.98.1.394

Jungraithmayr, Hermann. 1971. The Tangale vowel harmony system reconsidered. *Journal of African Languages* 10(1). 28–33.

Jungraithmayr, Herrmann. 1980. Kontakte zwischen Adamawa-Ubangi und Tschad-Sprachen: Zur Übertragung grammatischer Systeme. *Zeitschrift der deutschen morgenländischen Gesellschaft* 130(1). 70–85.

Jungraithmayr, Herrmann. 1981. Über die Mawa (Guera, Tschad) – Ethnographische und linguistische Notizen. In Inge Hofmann (ed.), *Festschrift zum 60. Geburtstag von P. Anton Vorbichler.* Wien: Afro-Pub. 47–70.

Jungraithmayr, Herrmann with the collaboration of Djimé N'Gaba Tchéré, Nakouma Touri, Michka Sachnine & Jean-Pierre Caprile. 1990. *Lexique mokilko: mokilko-français et français-mokilko (Guéra, Tchad).* (Sprache und Oralität in Afrika 6). Berlin: Dietrich Reimer.

Jungraithmayr, Herrmann & Abakar Adams. 1992. *Lexique migama: migama-français et français-migama (Guéra, Tchad), avec une introduction grammaticale.* (Sprache und Oralität in Afrika 17). Berlin: Dietrich Reimer.

Jungraithmayr, Herrmann & Philibus Diyakal. 2013. Mushere, a Chadic language of five level tones. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 163. 297–307.

Jungraithmayr, Herrmann & Carsten Peust. 2019. Notes on Gàlì (Miltu), a moribund Chadic language (Republic of Chad). *Folia Orientalia* 56. 219–230.

Kuteva, Tania. 2001. *Auxiliation.* Oxford: Oxford University Press.

Leger, Rudolf. 1994. *Eine Grammatik der Kwami-Sprache (Nordostnigeria).* (Westafrikanistische Studien. Frankfurter Beiträge zur Sprach- und Kulturgeschichte Band 8). Köln: Rüdiger Köppe Verlag.

Lukas, Johannes. 1936. *Die Logone-Sprache im zentralen Sudan.* (Abhandlungen für die Kunde des Morgenlandes, 21/6: viii + 148). Leipzig: Deutsche Morgenländische Gesellschaft.

Lukas, Johannes. 1937. *Zentralsudanische Studien.* (Hansische Universität-Abhandlungen aus dem Gebiet der Auslandskunde. Bd. 45. Reihe B: Völkerkunde, Kulturgeschichte und Sprachen Band 24.) Hamburg: Friedrichsen, de Gruyter & Co.

Lukas, Johannes 1971. Die Personalia und das primare Verb im Bolanci (Nordnigerien). *Afrika und Übersee* 55. 114–139.

Mahamat, Adam. 2005. *Esquisse de la phonologie lexicale du mpádì (langue tchadique centrale: groupe B).* Yaoundé: Université de Yaoundé 1 Mémoire de DEA.

DOI 10.15460/auue.2025.98.1.394

Melis, Antonio. 1999. *Description du Masa (Tchad). Phonologie, syntaxe et dictionnaire encyclopedique.* Tours: University of Tours PhD dissertation.

Meyer-Bahlburg, Hilke. 1972. *Studien zur Morphologie und Syntax der Musgu.* Hamburg: Helmut Buske.

Newman, Paul. 2000. *The Hausa Language. An encyclopedic reference grammar.* New Haven: Yale University Press.

Newman, Roxana Ma. 1971. *A case grammar of Ga'anda.* Los Angeles: UCLA Ph.D. dissertation.

Nougayrol, Pierre. 1999. *Les parlers gula (République centrafricaine, Soudan, Tchad). Éléments de grammaire comparative et lexique.* Paris: CNRS.

Pearce, Mary & Joseph Lovestrand. 2024. Vowel harmony in Chadic languages. In Ritter, Nancy & Harry van der Hulst (eds.), *Oxford handbook of vowel harmony.* Oxford: Oxford University Press. 603–609.

Persson, Andrew M. & Janet R. Persson 1991. *Mödö-English Dictionary with Grammar.* (Bilingual Dictionaries of Sudan 1). Nairobi: Summer Institute of Linguistics-Sudan.

Pohlig, James N. 1992. *An account of Mofu-Gudur verb infixes and suffixes.* Ministry of Scientific and Technical Research. Yaoundé: SIL.

Schneeberg, Nan. 1971. Sayanci verb tonology. *Journal of African Languages (Special Chadic Issue)* 10(1). 87–100.

Schuh, Russell. 2001. "Miya as a West Chadic language with V ... S word order." In Dymitr Ibriszimow, Rudolf Leger & Uwe Seibert (eds.), *Von Ägypten zum Tschadsee. Eine linguistische Reise durch Afrika. Festschrift für Herrmann Jungraithmayr zum 65. Geburtstag.* Würzburg: Deutsche Morgenländische Gesellschaft (Ergon Verlag). 435–49.

Schuh, Russell. 2017. *A Chadic cornucopia.* Oakland: California Digital Library.

Shay, Erin. 1999. *A grammar of East Dangla: The simple sentence.* Boulder: University of Colorado Ph.D. dissertation.

Skinner, Margaret Gardner. 1979. *Aspects of Pa'anci grammar.* Madison: University of Wisconsin Ph.D. Dissertation.

Smith, Tony. 2010. *Phonology of Muyang.* SIL.

Tourneux, Henry. 1978. Première note sur le mbara. *Africana Marburgensia* 11(2). 27–32.

Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic analyses: The non-Bantu languages of North-Eastern Africa.* London: Oxford University Press.

DOI 10.15460/auue.2025.98.1.394

Viljoen, Melanie 2013. *A grammatical description of the Buwal language.* Melbourne: LA Trobe University Ph.D. dissertation.

Wolff, H. Ekkehard. 1983. *A grammar of the Lamang language (Gwa'd Lamang)*. (Afrikanistische Forschungen 10). Glückstadt: J. J. Augustin.

Wolff, H. Ekkehard & Ludwig Gerhardt. 1977. Interferenzen zwischen Benue-Congo und Tschadsprachen. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, Suppl. III.2. 1518–1542.

Ziegelmeyer, Georg. 2013. On the verbal system of Gashua Bade. *Afrikanistik-Aegyptologie-Online, Vol. 2014*. https://dipp.archiv.hbz-nrw.de/journals/https://www.afrikanistik-aegyptologie-online.de/archiv/2014/3756/

DOI 10.15460/auue.2025.98.1.394

Palatalization and labialization in the Chadic languages of Chad

James Roberts

SIL Chad
james_s_roberts@sil.org

Abstract
This paper[1] examines verb data from five Chadic languages of Chad, in order to show the extent to which the prosodies of labialization (LAB) and palatalization (PAL) are relevant to the synchronic analysis of their phonologies. Previous studies have established the synchronic functioning of prosodies in Chadic languages of the Central branch, but not for other branches of the family. Of the languages considered in this paper, one (Heɗe) is from the Masa branch, the other four from the Eastern branch: Somrai, Gabri, and Kabalay from the Chari-Logone subbranch, and Mawa from the Guéra subbranch. In Heɗe and Somrai, the vowel system can be reduced to a single underlying /a/, similar to the analysis claimed for many Central Chadic languages. Epenthetic [ɨ] is added to separate between consonants, and the distinctive prosodic features of LAB and PAL blanket whole words, creating a full set of surface vowels. A consequence of such a system is that high vowels like [ɨ] are marginal, while only the nonhigh vowel /a/ is fully phonemic. Gabri displays these characteristics, but to quite a lesser degree; in this language, PAL also seems to play a limited morphological role. The vowel system of Kabalay again shows a dichotomy in behavior between high and nonhigh vowels, but there is only a hint that prosodies are relevant. Mawa clearly shows that the vowel /a/ can be affected by LAB and PAL, but these prosodies do not interact with the seven other phonemic vowels. The article concludes by suggesting that the LAB and PAL prosodies must have functioned in Proto-Chadic, although they have been lost to differing degrees in individual languages in the Chadic languages of Chad.

Keywords: East Chadic, Chad, prosodies, labialization, palatalization

1 This article is a revised version of the paper presented at the 4th Biennial International Colloquium on the Chadic Languages, Bayreuth, October 2007, entitled “Palatalization and Labialization in Eastern Chadic verbs.”

DOI 10.15460/auue.2025.98.1.392

1 Introduction

Chadic languages have a number of original features of interest to phonologists on both a descriptive and theoretical level. Perhaps the most intriguing phonological phenomenon concerns the suprasegmental features of palatalization and labialization. These two prosodies are especially prominent in the Central branch of Chadic, and their functioning has been described in a number of papers, e.g. Barreteau 1987, Roberts 2001, Wolff 2003, 2023. This phonological phenomenon is included in Schuh's (2017) massive Chadic compendium; in addition, he surveys the use of these prosodies with a morphological function. Schuh himself had described the morphological function of palatalization in West Chadic languages (Schuh 2002). But to my knowledge there are no clear references to these prosodies functioning in languages of the Eastern or Masa branches of the Chadic family. The burden of this paper is to demonstrate that prosodies are indeed relevant to the phonologies and morphologies of a number of Chadic languages in Chad. Admittedly, the prosodies are not of universal relevance in the Eastern or Masa branches, as they are in the Central branch. Nonetheless, I will look briefly at five languages and the synchronic evidence they contain for the functioning of palatalization (PAL) and labialization (LAB) as suprasegmental prosodies within their phonological systems. These prosodies are distinctive elements of the phonology, suprasegmental features which characterize whole words. Their effect is always seen in the vowels, but they may affect certain consonants as well. I conclude by considering the implications for our understanding of Proto-Chadic and the nature of its phonological system.

The data used in this paper will be limited to the verbs of these languages. But why verbs? In my experience, the morphology and phonology of verbs in Eastern Chadic languages is very tightly constrained and usually almost completely regular. Verbs borrowed from another language will be made to fit the phonological constraints imposed by the language. The same is not true of nouns, whose structures are much more disparate and open to diverse influences. As a result, any important patterns in the language are likely to be more clearly seen in the verbal system than anywhere else. Chadic verbs are all constructed from a certain number of consonants (up to as

DOI 10.15460/auue.2025.98.1.392

many as six) and no more than one vowel (see Frajzyngier 1983: 125). It should be noted that monoconsonantal verbs do display some peculiarities: this was the theme of an earlier Chadic colloquium (Jungraithmayr & Tourneux 1990). Apart from exceptional cases such as these, all verbs of these languages adhere to a single set of principles, which are specific to the given language.[2]

For the purposes of this paper, I have chosen five languages of Chad that show evidence of PAL and LAB operating to varying degrees. One language is taken from the Masa branch, and the others are Eastern Chadic: of these, three are from the Chari-Logone subgroup, and the last one is from the Guéra subgroup.

The first language is Heďe (ISO [hed]), a language of the Masa branch. The verb data on this language comes from Vaïbra (2003), who focuses on the phonology and morphology of the verbal system; he uses a corpus of 445 verbs. The next three languages are from the Chari-Logone subgroup of Eastern Chadic (Group A in Newman's (1977) classification of East Chadic): Somrai, Gabri, and Kabalay. For Somrai (ISO [sor]), SIL linguist Thomas Deusch collected and verified 260 verbs for use in a sketch of the phonological system of the language (Deusch 2007); a full analysis of prosodies in the Somrai verbal system is detailed in Roberts 2012. The Kabalay (ISO [kvf]) data come from a corpus of 320 verbs collected and treated more fully in a mémoire describing the verbal system of that language (Selgué 2005). For Gabri (ISO [gab]), I rely on the data collected by two native speakers collected over a period of several years, and produced in manuscript form as Kaïndi, Roberts & Samane 2000. This lexicon includes 802 basic verbs, not including derived verbs. It should be noted that this is the Darbé variety of (southern) Gabri, which is somewhat different from northern Gabri (ISO [tng]), properly called Tobanga and treated in Caprile 1978. Finally, Mawa (ISO [mcw]) represents the Guéra subgroup of Eastern Chadic. The corpus of 361 verbs was collected and checked by myself in the village of Mahoua with native speakers Youssouf Hissène and his father Hissène Abdoullaye; the essence of the Mawa analysis is presented in Roberts 2009. The sources of data for each language is summarized in Table 1.

2 I have not found synchronic evidence of verb classes defined by the final vowel and tone in Eastern Chadic, as claimed by Newman (1975) at least for the Western branch.

DOI 10.15460/auue.2025.98.1.392

Table 1. Languages and sources of data

Language	Source	Corpus
Heɗe	Vaïbra 2003	445 verbs
Somrai	Deusch 2007, Roberts 2005	260 verbs
Gabri	Kaïndi, Roberts & Samane 2000	802 verbs
Kabalay	Selgue 2005	320 verbs
Mawa	Roberts 2009	361 verbs

2 Heɗe

I start with Heɗe (ISO [hed]), one of the languages of the Zime cluster, belonging to the Masa branch and spoken in the area around the town of Pala.[3] As far as prosodies are concerned, this language behaves very much like the Central Chadic languages spoken across the border in Cameroon, so my presentation of this language will serve to illustrate how a language with fully functioning prosodies operates. I will not present arguments for the prosodic approach here; the aforementioned sources (Barreteau 1987, Roberts 2001, and Wolff 2003, 2023) develop this analysis in detail, and show why it is the most appropriate treatment of the phonology of Central Chadic languages. I will simply show here how the approach is applied to Heɗe.

Heɗe has six surface vowels which divide into two groups: the high vowels [i, ɨ, u] and the nonhigh vowels [e, a, o]. The vowel [ɨ] has a very limited distribution, especially in the verbal system. It is never the only vowel in a word; furthermore, it only occurs in a light syllable followed by a syllable containing [a]. It is best treated as an epenthetic vowel that is inserted when certain conditions are met. Consider now the following verbs, taken from Vaïbra 2003, shown in their base form (which is used both as verbal noun and as a finite incompletive form),[4] presented in Table 2:

3 The vùn dzìpàw variety described by Sachnine (1982) is considered to be a dialect of this language.

4 Transcriptions in this article follow IPA conventions, with an exception common to Africanists: [y] represents the palatal glide, and [j] the palatal plosive/affricate.

DOI 10.15460/auue.2025.98.1.392

Table 2. Heɗe verbs sorted according to root vowel and root structure

V	CV	CVC	CVCV	CVCVC
a	*pà* ‘love’	*ɬàt* ‘shake’	*sīɗā* ‘grow’	*kìràk* ‘trim’
e	*vè* ‘take’	*kēɓ* ‘look for’	*dîdē* ‘get wet’	*dîŋèr* ‘balance’
i	*tī* ‘eat’	*sīn* ‘send’	*wīlī* ‘shine’	*tīwīr* ‘turn’
o	*ŋgō* ‘control’	*kòl* ‘wait for’	*pūnō* ‘eat’	*hūrōk* ‘scratch’
u	*ɓū* ‘trick’	*cùm* ‘stretch’	*hūtū* ‘grind’	*tūpūr* ‘round off’

In this system, the vowel /a/ can be considered the only underlying vowel, if we admit of the two distinctive word-level prosodies of palatalization and labialization. The surface vowel [e] is seen as the prosody of palatalization realized on the vowel /a/, and [o] is treated as that same underlying /a/ when subjected to labialization. The situation is parallel among the high vowels: the surface vowel [i] is the realization of palatalization on the (otherwise epenthetic) “zero vowel” [ɨ], and [u] is the realization of labialization on that same vowel.[5] Along these lines, the underlying structure of the Heɗe verbs of Table 2 (apart from tone) can be represented as shown in Table 3. In this table, and in the rest of this article, I will use autosegmental representations to highlight the suprasegmental nature of PAL and LAB by according to these features a tier of their own, because they behave independently of the consonants and vowels. This practice is a logical extension of the insights of John Goldsmith and others (see Leben 2018), an approach which has even been exploited in treatments of Afroasiatic morphology (McCarthy 1981). For reasons of economy, I place all segments (consonants and vowels) on the same tier. The ∅ represents the absence of an underlying vowel; it will be filled by the insertion of the default vowel [ɨ]. When prosodies are also present, they attach to these two vowels to produce all of the surface vowels found in the data.

5 Note again the phonological weakness of the vowel [ɨ]: it must always appear in conjunction with another phonological element – either with one of the prosodies of palatalization or labialization, or else with the vowel /a/ in a separate syllable.

DOI 10.15460/auue.2025.98.1.392

Table 3. Underlying structure of Heɗe verbs

CV	CVC	CVCV	CVCVC
pà pa 'love'	*ɬàt* ɬat 'shake'	*sīɗā* sØɗa 'grow'	*kɨ̀rà̀k* kØrak 'trim'
PAL | *vè* va 'take'	PAL | *kēɓ* kaɓ 'look for'	PAL /\ *dīɗē* ɗØɗa 'get wet'	PAL /\ *dɪ̀ŋèr* ɗØŋar 'balance'
PAL | *tī* tØ 'eat'	PAL | *sīn* sØn 'send'	PAL /\ *wīlī* wØlØ 'shine'	PAL /\ *tīwīr* tØwØr 'turn'
LAB | *ŋgō* ŋga 'control'	LAB | *kòl* kal 'wait for'	LAB /\ *pūnō* pØna 'eat'	LAB /\ *hūrōk* hØrak 'scratch'
LAB | *ɓū* ɓØ 'trick'	LAB | *cùm* cØm 'stretch'	LAB /\ *hūtū* hØtØ 'grind'	LAB /\ *tūpūr* tØpØr 'round off'

In dissyllabic verbs (structure CVCV and CVCVC), it will be noted that the vowel of the first syllable is always high, and agrees in backness and roundness with the vowel of the second syllable: [i] with the front vowel [e], [ɨ] with the central [a], and [u] with the round [o].

The effect of these prosodies is not just limited to an individual vowel or syllable, but rather is applied to the word as a whole. This fact explains the vowel agreement phenomena alluded to earlier, and is confirmed when we consider the result of suffixing the object pronouns *-an* (1SG) and *-m* (3SG.M). Table 4 displays the results of adding these suffixes to verb roots of shape CVC.

DOI 10.15460/auue.2025.98.1.392

Table 4. Suffixation of *-an* (1SG) and *-m* (3SG.M) onto CVC verbs

Root V	CVC	Gloss	CVC + *-an*	CVC + *-m*
a	*ɬàt*	'shake'	*ɬìtàn*	*ɬìtìm*
e	*kēɓ*	'look for'	*kīɓēn*	*kīɓīm*
i	*sīn*	'send'	*sīnēn*	*sīnīm*
o	*kòl*	'wait for'	*kùlòn*	*kùlùm*
u	*cùm*	'stretch'	*cùmòn*	*cùmùm*

In the suffixed forms, the vowel of the root is deleted as a consequence of the morphological process of suffixation. Then the epenthetic zero vowel [ɨ] is inserted to separate the two consonants that have come into contact.[6] The shape of the suffix is not changed during affixation, although an additional [ɨ] will be inserted to separate the [m] of the 3SG.M suffix from the final consonant of the root. Note that if the verb root has a prosody, the prosody remains, and extends to affect the suffix as well as any epenthetic vowels, so that the whole word is covered by that prosody. The autosegmental representations of Figure 1 illustrate the process for the forms of *kēɓ* 'look for' and *cùm* 'stretch' (taken from Table 4).

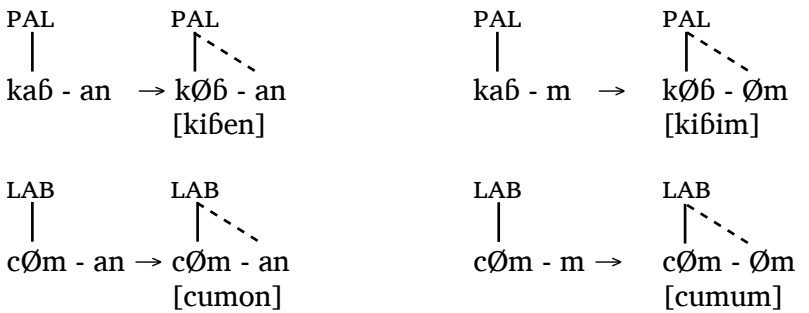


Figure 1. Effect of prosodies in inflected verbal forms of Heɗe

The consequence is that a prosody always applies to the word in its entirety, and not just the root.[7]

6 This description evokes the principles of Lexical Phonology, following Mohanan 1986. The deletion rule is a lexical rule triggered by morphological affixation; and the insertion of [ɨ] is undoubtedly a postlexical rule.

7 It should also be noted that in the inflected forms, there is neutralization of the vowel patterns for verbs with either [e] or [i] as root vowel, and also for verbs with either [o] or [u] as root vowel.

DOI 10.15460/auue.2025.98.1.392

3 Somrai

The second language is Somrai, a member of the Chari-Logone group of Eastern Chadic. Somrai demonstrates the operation of the palatalization and labialization prosodies to the greatest degree of any of the Eastern branch languages that I have examined. The verbal system of this language was described by Barreteau & Jungraithmayr (1979), who bring out much of the data relevant to the prosodies. However, they do not recognize the prosodies as such, and our interpretation of the data differs in a number of points from theirs. The verb data that I will be using, taken from Deusch (2007) and synthesized in Roberts (2005, 2012), is shown in Table 5.

Table 5. Inflected forms of Somrai verbs, sorted according to root vowel

Root vowel	Verbal noun	Perfective	Subjunctive	Gloss
a	*dàm-à*	*dàm*	*dàm-nà*	‘stay’
	ɓābīr-ā	*ɓābīr*	*ɓābīr-nà*	‘hang’
e	*ɗ̪èàs-à*	*dɛ̀ʃ(ì)*	*ɗ̪èàsɨ̀-nà*	‘close’
	ɗ̪ēāŋs-ā	*dɛ̄ŋsɪ̄*	*ɗ̪ēāŋsɪ̄-nà*	‘tamp’
o	*cu̱ār-á*	*cōŕ*	*cu̱āŕ-nà*	‘return’
	ju̱āŋg-á	*jōŋgɨ́*	*ju̱āŋgɨ́-nà*	‘stoop’
ɨ	*pɨ́d-ə̄*	*pɨ́d(ɨ́)*	*pɨ́dɨ́-nə̀*	‘pass’
	dɨ̀rs-ə̀	*dɨ̀rsɨ̀*	*dɨ̀rsɨ̀-nə̀*	‘flatten’
i	*wīd-ə̄*	*wīd(ɪ̄)*	*wīdɪ̄-nə̀*	‘get drunk’
	ʃɪ̄ʃm-ə̄	*ʃɪsɪ̄m*	*ʃɪsɪ̄m-nə̀*	‘suck’
u	*cūb-ə̄*	*cūb(ɪ̄)*	*cūbɪ̄-nə̀*	‘take off’
	ɓūms-ə̄	*ɓūmsɪ̄*	*ɓūmsɪ̄-nə̀*	‘germinate’

The verbs in this table are representative of the whole system. For each of the six surface vowels which occur in verb roots (seen most transparently in the perfective form), two verbs are displayed, one having two consonants and another with three consonants. The lexical vowel always occurs in the first syllable. After that, an epenthetic [ɨ] may be inserted predictably, in either of two contexts: (a) in word-final position of the perfective form, to support a final obstruent consonant; or (b) word-internally, to separate two adjacent consonants according to the principles of sonority sequencing, which I will not detail here. There is a major dichotomy between the behavior of the

DOI 10.15460/auue.2025.98.1.392

high vowels [i], [u], and [ɨ] on the one hand, and of the nonhigh vowels [e], [o], and [a] on the other, shown by the heavy line across the middle of the table. A verb with a nonhigh lexical vowel takes suffixes with the low vowel [a], while a verb with a high vowel has[ə] in its suffixes instead. What is more, the nonhigh vowels [e] and [o] show variability in their realization. In the morphologically bare perfective form, these vowels are monophthongs, yet in the subjunctive and nominal forms they show up as the corresponding diphthongs [e̯a] and [u̯a].

This whole array of data can be interpreted as the outworking of a simple underlying system which admits of only one underlying vowel, namely /a/, in conjunction with the two prosodies of palatalization and labialization, just as presented above for the Heɗe system. The verbs in the first six rows of Table 5 have the underlying vowel /a/, whereas the verbs in the last six rows have no underlying vowel at all; the surface vowels that appear there are the reflexes of the epenthetic vowel [ɨ] which is inserted to break up the consonants. The prosody of palatalization is present in those verbs with the surface vowels [e] and [i]. It fuses completely with the underlying /a/ to produce [e] in the perfective form, but fuses only partially with /a/ in the nominal and subjunctive forms to create the diphthong [e̯a]. The surface vowel [i] is the manifestation of palatalization as it affects the epenthetic vowel [ɨ], and is uniform in its realization. Similarly, the prosody of labialization is present in verbs with surface vowels [o] and [u]. The vowel [o] seen in the perfective forms is the surface result of labialization as it fully affects the underlying vowel /a/, and the diphthong [u̯a] seen in the nominal and subjunctive forms is the result of the partial fusion of labialization with this same underlying /a/. When there is no underlying vowel, there is no variability in the realization of the prosody of labialization: it appears as surface [u] in all forms of verbs that contain it. Finally, we note that verbs with surface vowels [a] and [ɨ] involve no prosody at all.

Table 6 presents my interpretation of the underlying phonological structure of each of the verb roots found in Table 5. A verb may have the vowel /a/, or else no underlying vowel at all (indicated here by Ø), in the first syllable. In addition, a verb may have no prosody at all, or else PAL, or else LAB. The surface forms of these roots (equivalent to the bare perfective forms) are shown to the side.

DOI 10.15460/auue.2025.98.1.392

Table 6. Underlying and surface structure of Somrai verbs

Root V	Underlying	Surface	Gloss
a	d a m	*dam*	‘stay’
	ɓ a b r	*ɓab(ɨ)r*	‘hang’
e	PAL d a s	*ɗeʃ*	‘close’
	PAL ɗ a ŋ s	*ɗeŋs(ɨ)*	‘tamp’
o	LAB c a r	*cor*	‘return’
	LAB j a ŋ g	*joŋg(ɨ)*	‘stoop’
ɨ	p Ø d	*pɨd(ɨ)*	‘pass’
	d Ø r s	*dɨr(ɨ)s*	‘flatten’
i	PAL w Ø d	*wid(ɨ)*	‘get drunk’
	PAL s Ø s m	*ʃɨs(ɨ)m*	‘suck’
u	LAB c Ø b	*cub(ɨ)*	‘take off’
	LAB ɓ Ø m s	*ɓums(ɨ)*	‘germinate’

The prosodies in Somrai work differently from Heɗe in that the prosody does not affect the vowels of every syllable of the word, but only and always the initial syllable. Nonetheless the prosodies are word-level features in that they may not affect vowels or syllables individually. A prosody is either present or absent for each verb root as a whole, as was the case in Heɗe.

There is an interesting sidelight concerning one of the consonants that confirms that the phenomenon involved here is suprasegmental, and not just a local segmental issue for vowels. When the consonant /s/ occurs in the first syllable of a verb with palatalization, it will be realized not as phonetic [s], but rather as the palatalized [ʃ], just as would happen in a Central Chadic language. The verb [ɗeʃ], for example, is underlyingly /PALɗas/, and [ʃisɨm] is underlyingly /PALs(ɨ)s(ɨ)m/. When the /s/ falls outside the influence of the prosody (i.e. in a noninitial syllable), it is realized as alveolar and not palatal: compare the occurrence of [ʃ] in the verbal noun form [ʃiʃ.mə] from Table 5 with the occurrence of [s] in the second syllable of the perfective form [ʃi.sɨm].

DOI 10.15460/auue.2025.98.1.392

4 Gabri

Gabri is another language of the Chari-Logone subgroup, spoken to the southwest of the Somrai-speaking area. It is possible to see the effect of prosodies in Gabri, although to a lesser degree than in Somrai. In fact, a traditional phonology of Gabri would not even be tempted to propose palatalization or labialization as relevant to its description. Nonetheless, one curiosity of Gabri phonology is that the high vowels [i] and [u] are only marginally phonemic, if they must be recognized at all: [i] only occurs in word-final position, where one might argue that it is merely a vocalized realization of the consonant [y]. Similarly, [u] occurs in only two positions: word-finally, where it could be treated as the vocalized realization of the consonant [w], and also in a syllable preceding another round vowel, where it could be reckoned as an [ɨ] which has undergone assimilation of rounding. The vowel [ɨ] is ubiquitous, but its function seems merely to provide transition between consonants that are not separated by any other vowel, or (at the end of words) to provide support to obstruents which occur word-finally. So this vowel also may be seen as epenthetic, and not an underlying vowel phoneme.

In the verbal system, we find that only very few vowels occur. The only vowels found in verb roots are the nonhigh vowels [a] and [ɔ/o], as well as the epenthetic zero vowel [ɨ]. There is also a handful of verbs with [u] (0.4% of my corpus, or 29 verbs), but in these verbs the [u] is either in absolute initial position, or else it follows the velar consonants, [k] or [g].[8] Table 7 displays the possibilities, with four inflectional forms given for each verb; suffixes are separated from the stem by a hyphen:

8 The evidence of contrast between [u] and the other vowels relies uniquely on these 29 verbs; for this reason the phonemic status of /u/ is extremely marginal within the language as a whole. Regarding its co-occurrence with [k] and [g], there is a well-known affinity between vowel-rounding, or labialization, with the velar consonants in Central Chadic languages (Barreteau 1987: 172, Roberts 2001: 95–96).

DOI 10.15460/auue.2025.98.1.392

Table 7. Inflected forms of Gabri verbs, sorted by root vowel

Root vowel	Verbal noun	Perfective	Imperative[9]	Ventive[10]	gloss
a	*làs-è*	*làs(ī)*	*làs-à*	*lès-è*	'encourage'
	ɓágɪ́m-ē	*ɓágɪ̄m*	*ɓágɪ́m-ā*	*ɓégɪ́m-ē*	'pick up'
	hìràg-è	*hìràg(ī)*	*hìràg-à*	*hìrèg-è*	'bemoan'
o/ɔ	*bɔ̀s-ɛ̀*	*bòs*	*bìs-à*	*bìs-è*	'force'
	sɔ́bír-ɛ́	*sóbír*	*sɪ́bír-á*	*sɪ́bír-é*	'control'
	bùdɔ̀l-ɛ́	*bùdǒl*	*bɪ̀dɪ̀l-á*	*bɪ̀dɪ̀l-é*	'vomit'
ɪ	*dɪ̀b-é*	*dɪ̀b(ī)*	*dɪ̀b(ī)*	*dɪ̀b-è*	'put'
	mɪ̀gír-ē	*mɪ̀gɪ̄r*	*mɪ̀gír*	*mɪ̀gír-ē*	'crunch'
u	*gùd-ē*	*gùd(ī)*	*gùd(ī)*	*gùd-ē*	'knead'

This table, like the preceding Table 6 for Somrai, distinguishes by a heavy line between the nonhigh lexical vowels [a] and [o/ɔ] above the line, and the high vowels [ɪ] and [u] below. It will be noted that regardless of the number of consonants contained in the verb, the lexical vowel ([a] or [o/ɔ]) only occurs once in the word.[11] For the verbs of Table 7 which have three consonants, it will be noted that the position of that vowel is variable, occurring either in the first or in the second syllable. Another curiosity is the variation between [o] and [ɔ], wherein the open variety [ɔ] always occurs in the verbal noun form, and the closer variety [o] always occurs in the morphologically bare perfective form. Speakers systematically make the difference between these two vowels, although no minimal pairs can be produced to show a contrast. The same situation obtains for the pair of front vowels [e] and [ɛ].

These data do not make an irrefutable argument for an analysis with prosodies, but there are a number of elements that are remi-

9 The imperative might more properly be called the irrealis form, for it is used in a variety of contexts in Gabri, including the conditional, negative, etc.

10 The ventive form, although created productively, is not used very frequently in the language. Its principal usage is in the imperative (and other irrealis contexts), and it actually looks like it is based on the imperative form by a change of vowel(s). However, even verbs with root vowel [ɪ] or [u], which have no suffix in the imperative, still add a suffix *-e* in the ventive.

11 In common with all Eastern Chadic languages. Gabri allows only one lexical vowel per verb. Frajzyngier (1983: 125) concurs with this claim.

DOI 10.15460/auue.2025.98.1.392

niscent of prosodic systems found elsewhere. I suggest that the unusual distributional facts can be explained if an earlier stage of the language had a more fully functioning prosody system which has been largely eroded away today. The near non-existence of distinctive high vowels is of course one hallmark of the prosodic systems of Central Chadic, as well as of Heɗe and Somrai. A parallel to that situation is the similarity in behavior between the nonhigh vowels. In Gabri, while verbs whose vowel is [ɨ] or [u] take no suffix in the imperative form, verbs with either of the nonhigh lexical vowels [a] or [o/ɔ] take the suffix *-a.* Curiously, though, the *o*-verbs have no [o/ɔ] in the imperative form. If the *o*-verbs are simply considered to be *a*-verbs affected by the labialization prosody, it might be claimed that the prosody, along with the underlying /a/ vowel that it affects, drops out in the imperative.

If *o*-verbs do involve the prosody of labialization, the behavior of the prosody is a bit peculiar. The LAB prosody first docks onto the underlying vowel /a/ in the root, and then spreads backward from that point to the beginning of the word, transforming the epenthetic [ɨ] into [u]. Figure 2 shows an autosegmental representation of the *o*-verbs *budol* and *sobɨr*. In the latter verb, note that the vowel [ɨ] in the syllable following o is unaffected, because lab spreads only to the left.[12]

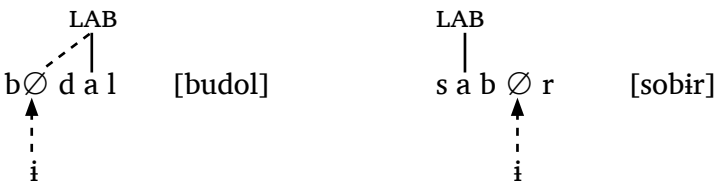


Figure 2. Effect of the LAB prosody in the *o*-verbs of Gabri

It is interesting to note that, although the lexical vowel of Gabri verbs may be [a] or [o], there are no verbs with a front vowel (especially the nonhigh [ɛ] or [e], nor even with the high vowel [i]). There may be a morphological reason for this. The ventive form, shown in Table 7, takes the verb root stripped of any prosody, and adds the

12 Barreteau (1987: 177) shows a similar peculiarity in the behavior of the LAB prosody in the Central Chadic language Mafa. In that language, LAB parses the word to be affected from the right edge, docking onto the first velar consonant it encounters, then affects all vowels and velar consonants further to the left from that point. If the word contains no velar consonant, LAB associates only to the leftmost vowel. Roberts (2001: 104) shows this in an autosegmental representation.

DOI 10.15460/auue.2025.98.1.392

suffix *-e*. At the same time, any instances of [a] within the root are fronted to [e]. This situation is reminiscent of similar phenomena in Podoko (Central Chadic; see Swackhamer n.d., Roberts 1994) or in Miya (West Chadic; see Schuh 2002), where applying the prosody of palatalization to a verb invests it with an inflectional value.

In Gabri, I propose that the ventive form is characterized by the PAL prosody. In addition to PAL, which blankets the whole word, there is a suffix which is underlyingly /-a/, but shows up as [e] because of the prosody. Figure 3 shows the formation of the ventive for the verbs [ɓagɨme] and [hɨrage].

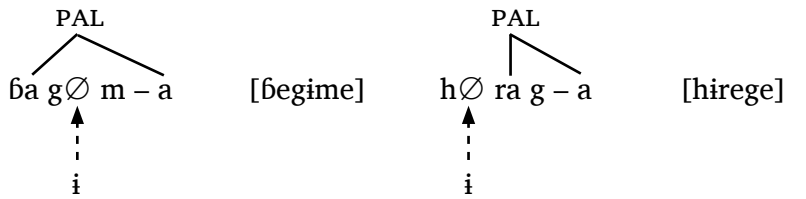


Figure 3. Formation of the ventive forms of Gabri

It will be noticed that PAL in these ventive forms does not affect the epenthetic [ɨ] by fronting it to [i], although in parallel fashion LAB does transform [ɨ] into [u] by leftward spreading, as seen in Figure 2. This fact simply demonstrates that the two prosodies operate independently and may have differing effects, just as is the case in their operation in Central Chadic languages (see Roberts 2001).[13] It also offers further confirmation that the evidence for LAB in Gabri is stronger than the evidence for PAL, and emphasizes the very marginal character of [i] in this language.

5 Kabalay

The next language to be considered is Kabalay, also taken from the Chari-Logone group, and the western neighbor of Gabri. The verb data here is very similar to that of Gabri. But if the case for prosodies in Gabri was weak, the case is even weaker in Kabalay. Nonetheless, there are features of this language which suggest that prosodies may

13 Alternatively, one might claim that there is no [i] in ventive forms because the epenthetic [ɨ] is added at a very late stage, after PAL has effected its changes. Even so, that analysis would have to recognize that PAL behaves differently from LAB.

DOI 10.15460/auue.2025.98.1.392

have been more active in the past, compared to what we see today. The Kabalay data under consideration are displayed in Table 8, taken from Selgue 2005:

Table 8. Verbal forms of Kabalay, sorted by root vowel

Root V	Verbal noun	Perfective	Imperative	Gloss
a	*bàr-á*	*bàr*	*bàr*	‘insult’
	ɲɨlāŋ-á	*ɲɨlāŋ*	*ɲɨlāŋ*	‘slip’
o	*bōl-á*	*bōl*	*bāl*	‘get pale’
	cōgīm-á	*cōgīm*	*cāgīm*	‘grab’
ə	*sə̄m-á*	*sə̄m*	*sīm-ā*	‘whip’
	cə̄gīn-á	*cə̄gīn*	*cīgīn-á*	‘sew’
ɨ	*sɨ̄l-á*	*sɨ̄l*	*sɨ̄l*	‘cut down’
	lɨ̄gɨ̄r-á	*lɨ̄gɨ̄r*	*lɨ̄gɨ̄r*	‘encircle’
u	*gūn-á*	*gūn*	*gūn*	‘poison’
	lūgīr-á	*lūgīr*	*lūgīr*	‘sort’
i[14]	*nín-á*	*nín*	*nín*	‘wander’

The three forms given for each verb in Table 8 correspond to the first three forms shown in Table 7 for the verbs of Gabri. One peculiarity of Kabalay is that the vowel [e] is virtually nonexistent; instead, we find the phonetic vowel [ə] in its place. I do not know why a former *[e] should have centralized globally in Kabalay, if that is the source of the vowel [ə] that we see today. At any rate, [ə] functions fully as a nonhigh vowel like [a] and [o].

Consider now the shape of the imperative forms. When the root vowel is high [i, u, ɨ], there is nothing surprising: the verbal noun takes the suffix *-ə*, and there is no suffix in either the perfective or imperative. The remarkable case is in the imperative form of verbs whose root has a nonhigh vowel [a, o, ə]. In those verbs, the imperative always has the vowel [a], regardless of which was the original lexical vowel. If the root vowel was [o], the [a] replaces it *in situ* in the imperative form. But if the root vowel was [ə], the vowel disappears altogether from the stem; the vowel [a] appears as a suffix. As elsewhere, the epenthetic [ɨ] separates contiguous consonants.

14 Verbs with a robust lexical vowel [i] are very rare. Selgue (2005) does not show any with three consonants, so there are none to show in Table 8, as for verbs with the five other lexical vowels.

DOI 10.15460/auue.2025.98.1.392

Here again we have only a vague vestige of the operation of prosodies. However, Kabalay still manifests a dichotomy in behavior between the high and nonhigh vowels, indicated again in Table 8 by the solid line that separates the two sets. Although the high vowel [u] is relatively well established as phonemic, [i] is very marginal, almost always occurring in the company of a palatal consonant. The nonhigh vowels, on the other hand, are all very robust as surface phonemes. With respect to the nonhigh vowels, the prosodic hypothesis would interpret [o] as an underlying /a/ affected by the LAB prosody, and [ə] would be an underlying /a/ affected by the PAL prosody (although a fronted realization is masked by a further process of centralization). Accordingly, one might claim that the imperative is formed by stripping off any prosody that might be present. And that provides at least a weak explanation for why all verbs with a lexical nonhigh vowel show the vowel [a] in the imperative, because they all have /a/ as an abstract underlying vowel.

6 Mawa

The final language to be considered is Mawa, the only one taken from the Guera group (Group B of Newman 1977) of East Chadic. Together with Ubi, Sokoro, Saba, and Barein, Mawa belongs to the southern subgroup. The verb data to be considered is based on my own field notes; these data are presented more fully in Roberts 2009, where I argue that distinctive prosodies of palatalization and labialization are the best means of accounting for certain phenomena in the language. I should also say that Mawa is the only language of the Guera group that I have found to show clear evidence for prosodies. Most of the other languages in this group have classic five-vowel inventories, and a prosodic analysis is not helpful to describe them.

Mawa is thus a very interesting case. The verb data to be considered are presented in Table 9; all verbs are shown in their infinitival, or verbal noun, form:

DOI 10.15460/auue.2025.98.1.392

Table 9. Verbal noun forms of Mawa verbs

Root vowel	Short vowel	Long vowel
a	*dàgāŋ* ‘lose’	*də̀àsāŋ* ‘jump’
o	*tòrōŋ* ‘bite’	*jòòpōŋ* ‘cover’
e	*tèpēŋ* ‘think’	*dɛ̀èmēŋ* ‘break’
ə	*làwəŋ* ‘insult’	*kə̀əlāŋ* ‘meet’
u	*dʊ̀lūŋ* ‘bow’	*sùùgūŋ* ‘move away’
i	*wìcīŋ* ‘swell’	*tɪ̀ɪ̀sīŋ* ‘fill a hole’
u̯a	*du̯ànāŋ* ‘lean on’	*lu̯ààkāŋ* ‘go around’
e̯a	*me̯àtāŋ* ‘groan’	*de̯ààsāŋ* ‘repair’

The suffix of the verbal noun form is *-ŋ*, which appears on all forms in this table. As is true for the other languages surveyed above, each verb has exactly one lexical vowel; in Mawa the lexical vowel fills all the vowel slots of the word. Unlike the preceding languages, however, the lexical vowel in a Mawa verb is maintained without change in all its inflectional forms, and for that reason the inflected forms have not been included in this table. Vowel length is distinctive in Mawa, but a long vowel may only occur in the first syllable. Each of the eight vowels of the above table is distinctive, occurring in a wide range of contexts.

The justification for prosodies in Mawa is restricted to the diphthongs [e̯a] and [u̯a] seen in the last two lines of Table 9. These diphthongs are phonetically identical to the ones that occur in Somrai. But in Mawa these diphthongs contrast with the simple vowels: in particular, [e̯a] contrasts with [e], and [u̯a] with [o]. Examples of these contrasts are given in (1):

(1) a. *e̯amaŋ* ‘push’
emeŋ ‘crunch’

b. *u̯aalaŋ* ‘cool off’
ooloŋ ‘hurt’

Even if Mawa once had a minimal vowel inventory like Hɛdɛ or Somrai, a full set of vowels has now become phonologized. Nonetheless, I argue (Roberts 2009) that its phonological analysis must still recognize the distinctive prosodies of PAL and LAB synchronically, even though they only affect words with the vowel /a/. The principal arguments in favor of word-level prosodies in Mawa rest on the

DOI 10.15460/auue.2025.98.1.392

suprasegmental behavior associated with the diphthongs. The prosodic approach provides an explanation for otherwise peculiar distributional facts. For one thing, there may only be one diphthong per word, and it will always occur on the initial syllable. Secondly, there is variability in the realization of the prosodies. The diphthong [e̯a] may alternatively be realized as the monophthong [ɛ], and [u̯a] may be realized as [ɔ]. In other words, a prosody which normally affects the vowel /a/ by breaking it into a diphthong, may optionally fuse with it to form a monophthong. Not only that, but other syllables beyond the first may sometimes be affected phonetically, as seen in the variability in words of (2). The autosegmental diagram of Figure 5 attempts to capture these facts. The solid association line shows where the prosody must be realized, and the dotted lines connect to the other segments that are optionally affected by it.

(2) a. [se̯atak] ~ [sɛtɛk] ~ [ʃɛtɛk] ‘hot pepper’
b. [ju̯akalam] ~ [jɔkɔlam] ~ [jɔkɔlɔm] ‘elbow’

Figure 5. Variability in the effect of PAL and LAB in Mawa

Further arguments for LAB and PAL in Mawa are presented in Roberts (2009), showing that the prosodic approach is superior to any alternative analysis. If prosodies are still functioning in Mawa today, it suggests that they are a vestige of system that once depended more heavily on such suprasegmental features. In this regard Mawa may be very conservative as compared to neighboring Chadic languages of the Guera, which have lost almost all evidence of a prosodic system such as that which is still very much active in the languages of Central Chadic.

7 Synthesis

What conclusions can be drawn from the data of these five languages? First, the facts from Heɗe and Somrai establish that full-fledged synchronic prosodic systems with LAB and PAL exist in the Masa and East branches of the Chadic family, and are not limited to Central branch. In both of these languages, the vowel system can be reduced to a single underlying vowel /a/. An epenthetic [ɨ] is added

DOI 10.15460/auue.2025.98.1.392

to supply a default vowel when required, and word-level prosodies of palatalization and labialization are superimposed to produce the full range of surface vowels.

In Somrai and in Mawa, the prosodies produce diphthongs [e̯a] and [u̯a] whose broken nature allows us to hear the prosodies affecting the underlying vowel /a/ without fully fusing with it. In both languages also, there is some degree of alternation between each diphthong and a corresponding monophthong, whether in a morphophonemic or in an optional phonetic context. Mawa is quite different from Somrai, nonetheless, in that it has developed a full set of phonemic vowels. The high vowels function fully in the language, just as much as the nonhigh vowels. This characteristic would seem to detract from considering Mawa as having a prosodic system. Yet the vowel /a/, whether alone or affected by LAB or PAL, is by far the most common vowel overall in Mawa, found in 47% of all verbs. In other languages with a prosodic system, such as Heďe or Somrai, about half of the words are analyzed as having an underlying /a/.

Prosodies play a much reduced role in Gabri, as compared to Somrai. Nonetheless, some of the same characteristics hold. For one thing, the status of high vowels in Gabri is very marginal; the nonhigh vowels [a] and [o], and potentially also [e], display a similarity of behavior in the imperative forms. In this language it is also possible that palatalization functions morphologically, and not simply phonologically. In Kabalay, it was established that all three nonhigh vowels are reduced to the basic vowel /a/ in the imperative form, as though any prosody which might have been lexically present in a verb is stripped off by a morphological process. The high vowels [i̯, u, i] are marginal in this language too, just as in other languages with a functioning prosodic system. Apart from those phenomena, though, a prosodic analysis does not seem very helpful in describing Kabalay.

Note that the five languages chosen for this study are not all spoken in the same geographical region. Gabri, Kabalay, and Somrai are all neighboring languages, it is true, but Heďe is a bit further afield, and there is quite a geographical and cultural separation between these four languages and Mawa. Nonetheless, synchronic analyses of all these languages show some relevance of palatalization and labialization to their phonological (and maybe also morphological) systems, to a greater or lesser degree. The simplest explanation for the existence of prosodic systems in scattered languages throughout

DOI 10.15460/auue.2025.98.1.392

the Eastern zone is to suppose that prosodies were characteristic of the Proto-language. And since prosodies still play such a fundamental role in languages of the Central branch, it might be supposed that prosodies were part of Proto-Chadic itself.

If Proto-Chadic did indeed make full use of prosodies, then such a system has seriously eroded in the Eastern branch. The synchronic and comparative studies attempted in this paper show that the prosodies of LAB and PAL have ceased their function to differing degrees in individual languages, from the relatively complete retention of prosodies in Somrai, to the remote vestiges of them in Kabalay. Further synchronic studies of individual Chadic languages may suggest possible paths for understanding how the prosodies have been lost, and the mechanisms by which each individual language has evolved.

References

Barreteau, Daniel. 1987. Du vocalisme en tchadique. In Daniel Barreteau (ed.), *Langues et cultures dans le bassin du lac Tchad.* Paris: ORSTOM. 161–191.

Barreteau, Daniel & Herrmann Jungraithmayr. 1979. Le verbe en sibine. In Herrmann Jungraithmayr (ed.), *The Chad languages in the Hamitosemitic-Nigritic border area.* Berlin: Reimer. 192–229.

Caprile, Jean-Pierre. 1978. Notes linguistiques sur le tobanga à partir d'un conte en cette langue. In Herrmann Jungraithmayr & Jean-Pierre Caprile (eds.), *Cinq textes tchadiques.* Berlin: Reimer. 121–175.

Deusch, Thomas. 2007. Soumraye phonology sketch. N'Djaména: SIL Chad. Ms.

Frajzyngier, Zygmunt. l983. The underlying form of the verb in Proto-Chadic. In Herrmann Jungraithmayr (ed.), *The Chad languages in the Hamitosemitic Nigritic border area.* Berlin: Reimer.123–143.

Jungraithmayr, Herrmann & Henry Tourneux (eds.). 1990. *Etudes tchadiques. Verbes monoradicaux, suivis d'une note sur la négation en haoussa: Actes de la XIIème réunion du Groupe d'études tchadiques LACITO-CNRS-Paris.* Paris: Geuthner.

Kaïndi, Etienne, James Roberts & Samane Dana. 2000. Lexique gabri. N'Djaména. Ms.

Leben, William. 2018. Autosegmental phonology. *Oxford research encyclopia of linguistics.* Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.365

DOI 10.15460/auue.2025.98.1.392

McCarthy. John. 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12. 373–413.

Mohanan, K.P. 1986. *The theory of Lexical Phonology*. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-3719-2

Newman, Paul. 1975. Proto-Chadic verb classes. *Folia Orientalia* 16. 65–84.

Newman, Paul. 1977. Chadic classification and reconstructions. *Afroasiatic Linguistics* 5/1. 1–42.

Roberts, James. 1994. Nontonal floating features as grammatical morphemes. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 38. 87–99. DOI: 10.31356/silwp.vol38.11

Roberts, James. 2001. Phonological features of Central Chadic languages. In Ngessimo M. Mutaka & Sammy B. Chumbow (eds.), *Research mate in African linguistics: Focus on Cameroon.* Köln: Rüdiger Köppe Verlag. 93–118.

Roberts, James. 2005. *Phonological structure of Soumraye verbs.* N'Djaména: SIL Chad. Ms.

Roberts, James. 2009. Palatalization and labialization in Mawa (Eastern Chadic). In Eva Rothmaler (ed.), *Topics in Chadic linguistics V: Comparative and descriptive studies.* Köln: Rüdiger Köppe Verlag. 129–140.

Roberts, James. 2012. Phonological structure of Somrai verbs. Paper presented at WOCAL 7, Buea (Cameroon). Ms.

Sachnine, Michka. 1982. *Le lame (vùn dzəpàò): un parler zimé du Nord-Cameroun (langue tchadique)*. Paris: SELAF.

Schuh, Russell. 2002. Palatalization in West Chadic. *Studies in African Linguistics* 31(1/2). 97–128. https://doi.org/10.32473/sal.v31i1.107352

Schuh, Russell. 2017. *A Chadic cornucopia. UCLA: e-press.* https://escholarship.org/uc/item/5zx6z32d

Selgue, Mahamat. 2005. *La description du système verbal du kabalaye.* N'Djaména: Université de N'Djaména mémoire de maîtrise.

Swackhamer, Jeanette. n.d. Podoko phonology. Yaoundé: SIL Cameroon. Ms.

Vaïbra, Nicolas. 2003. *Essai de description du système verbal de « vùn hédē » (langue zìmè)*. N'Djaména: Université de N'Djaména mémoire de maîtrise.

Wolff, H. Ekkehard. 2003. Segments and prosodies in Lamang and Hdi. Paper presented at WOCAL 4 17–22 June 2003 at Rutgers University in New Brunswick.

DOI 10.15460/auue.2025.98.1.392

Wolff, H. Ekkehard. 2023. *Lexical reconstruction in Central Chadic: a comparative study of vowels, consonants and prosodies.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009346375

DOI 10.15460/auue.2025.98.1.392

The vowel system of Ndam

James Roberts

SIL Chad
james_s_roberts@sil.org

Abstract
Ndam (Eastern Chadic, ISO [ndm]) displays an array of seven or nine surface vowels. However, the distribution and behavior of these vowels, as evidenced from morphophonemic data, shows that the inventory can be reduced to two basic vowels, /ə/ and /a/, if the operation of two prosodies is also admitted. These prosodies, labialization and palatalization, are distinctive units in the phonology of the language, as well as the two basic vowels; together, they account for the full range of surface vowels and their alternations, as shown in this article. The prosodic analysis, a hallmark of the analysis of Central Chadic languages, provides another example of an Eastern Chadic language which exploits prosodies in its phonology (see Roberts 2009). The article concludes by claiming that Chadic vowel systems are all rectangular, with a fundamental dichotomy between high and non-high vowels. The conclusion also warns of the dangers of confusing the two vowels [ɨ] and [ə] in the transcription of Chadic languages.

Keywords: Eastern Chadic, prosodies, palatalization, labialization

1 Introduction

The nature of the underlying vowel systems of Chadic languages has been under scrutiny for a long time. Despite a plethora of surface vowels, numbering as many as eight, nine, or more, most analysts have found it possible and desirable to reduce the number of underlying vowels in these languages to a bare minimum, and in the most extreme case, to a single underlying vowel /a/. Most recently, Wolff (2024, forthcoming) has proposed that the underlying vowel system of Proto-Chadic was indeed of such a minimal nature, and that a set of prosodies was responsible for producing the array of surface vowels in present-day Chadic languages. Synchronic analyses

DOI 10.15460/auue.2024.97.1.387

of this sort have become commonplace for languages of the Central Chadic branch (Wolff 1981, Barreteau 1987, Roberts 2001, etc.), but such approaches have not been at all consensual for the Western, Eastern, and Masa branches of the family. Eastern Chadic languages of the Guéra group, for example, almost all display robust "classic" five-vowel systems which are reckoned to be typologically the most common vowel systems around the world (Maddieson 1984). Nonetheless, Roberts (2025) has pointed out vestiges of a minimal vowel system in a few of these Chadic languages: Somrai, Kabalay, Gabri from the Eastern branch, and Herdé (Zimé Pala) from the Masa branch. The present article adduces evidence for a minimal vowel system in Ndam, another Eastern Chadic language, and proposes an integrated perspective for the vowel systems of all Chadic languages.

2 Surface phonology of Ndam

Ndam (ISO [ndm]) is spoken in the East Tandjilé region of Chad, on the south bank of the Chari river, by at least 10.000 speakers. It has been classified in the Chari-Logone branch (group A of Newman 2013) of the Eastern Chadic family, and has not received much attention from linguists. One simple monograph by Michael Broß exists for the Dik dialect, dating from 1988. More recently, a missionary-linguist William Cray has produced a limited amount of data and field notes on the language (2012a, 2012b). I was also privileged to work briefly with a team of four Ndam speakers in 2021 and 2022[1] in order to hammer out a practical writing system for the language. Several hundred words and a few morphological paradigms were collected, and two stories were recorded and transcribed. Such are the data on which the present analysis is based.

An initial surface inventory of the consonants[2] of the language can be presented as in (1), and that of the vowels in (2). Segments in parentheses are clearly non-phonemic.

1 On both occasions, we worked in the compound of the Association Tchadienne d'Alphabétisation, Linguistique et Traduction de la Bible in Moundou, Chad. The four Ndam speakers were Abderamane Kadi Djimet, Ahmat Saleh, Nadjara Sabour, and Younous Saleh.

2 Transcriptions in this article follow the conventions of the International Phonetic Alphabet (IPA), with two exceptions: [j] represents the palatal plosive/affricate, and [y] the palatal glide.

DOI 10.15460/auue.2025.98.1.387

(1)	p	t	c	k	(ʔ)
	b	d	j	g	
	ɓ	ɗ	ʃ		
		s	(ʃ)		h
	m	n	ɲ	ŋ	
		l			
		r			
	w		y		

(2)	i	ɨ	u
	e	ə	o
		a	

The four native speakers of the language that I worked with reckon that each of these seven vowels needs to be retained in a practical alphabet for the language. In addition to these simple vowels, there are two diphthongs which occur with some regularity in the various transcriptions: [ya]/[ye] and [wa]. It may be worth noting that Boyeldieu (1985), in his analysis of the neighboring Niger-Congo language Niellim, posits these same two diphthongs as unit vowels; they also occur in another neighboring language, the isolate Laal, where Lionnet (forthcoming) again treats them as units. In any case, these diphthongs will prove to function on a par with the simple vowels in Ndam, so I add them to the vowel chart, shown in (3):

(3)	i	ɨ	u
	e	ə	o
	ya	a	wa

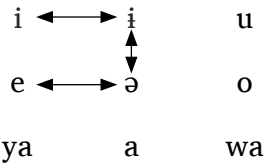
Vowel length is probably not distinctive in Ndam. In each of my three data sources (Broß 1988; Cray 2012a, 2012b; and my own research), there is a very limited number of lexemes transcribed with a long vowel, but that vowel is always [aa]. Further research will be needed to determine the exact status of such occurrences.

In addition, I note that there are a number of discrepancies in the transcription of certain vowel qualities across the three sources of my data, and an attempt to reconcile them constitutes a significant part of the present analysis. Specifically, these inconsistencies are in

DOI 10.15460/auue.2025.98.1.387

the variation between [ɨ] and [ə], and also between [i] and [ɨ], and finally between [e] and [ə], shown graphically in (4):

(4)



Apart from any discrepancies in transcription, though, there are several phonetic influences that pull a vowel's quality in different directions. The front vowels [i] and [e], for example, tend to centralize to [ɨ] and [ə], respectively, when they occur in closed syllables; examples of this are shown in (5).

(5) *ɓɨg* 'to pour', *ɨs* 'to cover', *kɨl* 'belly'
ɗəs 'to sneeze', *pəl* 'to breathe', *əg* 'to pound'

In fact, the closed (C)VC structure is the most common for Ndam lexemes. As a result, the front vowels may only retain their front position phonetically in open syllables, especially at the end of a word, shown in (6).

(6) *pi* 'to marry', *ni* 'to fail', *li.gər* 'to roll up'
ce 'to cut', *pe* 'to measure', *ma.le* 'to weed'

A counterbalance to the centralizing influence seen in the closed structures of (5) is the fronting influence of the palatal consonants [c, j, ʃ, ɲ, y]. These consonants occur more frequently next to front vowels [i] and [e] than to central vowels [ɨ] and [ə]. The palatals more commonly act regressively to affect a preceding vowel (rather than a following vowel), as we see in (7).

(7) *piye* 'to pour', *ɓireɲ* 'to break the neck', *dijɨm* 'to suck'

In contrast, the back vowels do not undergo phonetic influences, as a rule. Nonetheless, there is one interesting effect: the labial-velar glide [w] is never followed by the central vowel [ɨ]. In some cases it is clear that what should have been the vowel [ɨ] has been rounded by [w] to [u], as in (8):

(8) *wɨra* → [wura] 'walked (PST)'

DOI 10.15460/auue.2025.98.1.387

3 Evidence from morphophonemic alternations

Despite the potential confusion produced by all these influences, it is possible to determine the actual identity of each vowel and its relationship to the others when one considers the behavior of Ndam vowels in regular morpho-phonological alternations, of the sort that Jungraithmayr (1977) would call apophony. In my data, there are two cases of morpho-phonological alternation that are particularly helpful, one from the nominal system, and the other from the verb system. Both sets of morpho-phonological data show the same vowel alternations.

The markers of inalienable possession are suffixed directly onto the noun root, and display nine distinctions of person, in common with most Eastern Chadic languages. The paradigms of Table 1 show that the vowel [a] of the noun root changes to [ɨ] when the possessive suffixes are added:

Table 1. Inalienable possessive forms of Ndam nouns with /a/

	bad	‘hand’	*dagil*	‘foot’
1sm	*bɨd-an*	‘my hand’	*dɨgil-an*	‘my foot’
2sm	*bɨd-am*	‘your (m.) hand’	*dɨgil-am*	‘your (m.) foot’
2sf	*bɨd-ay*	‘your (f.) hand’	*dɨgil-ay*	‘your (f.) foot’
3sm	*bɨd-aw*	‘his hand’	*dɨgil-aw*	‘his foot’
3sf	*bɨd-at*	‘her hand’	*dɨgil-at*	‘her foot’
1ex	*bɨd-anu*	‘our (ex.) hand’	*dɨgil-anu*	‘our (ex.) foot’
1in	*bɨd-anan*	‘our (in.) hand’	*dɨgil-anan*	‘our (in.) foot’
2pl	*bɨd-aŋen*	‘your (pl.) hand’	*dɨgil-aŋen*	‘your (pl.) foot’
3pl	*bɨd-ak*	‘their hand’	*dɨgil-ak*	‘their foot’

The second example of vowel alternation comes from the verb system. Verbs in Ndam display a modest amount of morphology, principally achieved through suffixation. The citation form of the verb is its base, used directly as the nominal “infinitive” form, or else as a conjugated form which is unmarked for tense and aspect. In common with all Eastern Chadic languages, each Ndam verb has just one lexical vowel; Frajzyngier (1983: 125) takes the same position. Inflected forms involve suffixation onto the base, and in certain cases, a change in the root vowel accompanies the addition of a suffix. This is the case of the past tense form, which is compared with the base

DOI 10.15460/auue.2025.98.1.387

form in Table 2: in the past, the suffix *-a* is added to the base, but the lexical vowel [a] is also changed to [ɨ]:

Table 2. Base form and past form of verbs with vowel [a]

Base form of verb	Past form of verb
kal 'see'	*(in) kɨl-a* '(I) saw'
gad 'scratch'	*(in) gɨd-a* '(I) scratched'

This alternation of [a] with [ɨ] is identical to the alternation observed in Table 1. In the rest of this paper I will examine principally the verb data such as that in Table 2, and most notably verbs of structure CVC, the most common lexeme shape in the language.

The next alternations concern the two other central vowels of (4). The verbs of Table 3 have [ɨ] in their base form, and those in Table 4 have base vowel [ə].

Table 3. Base form and past form of verbs with vowel [ɨ]

Base form of verb	Past form of verb
gɨd 'bite'	*gɨd-ə*
lɨm 'braid'	*lɨm-ə*

Table 4. Base form and past form of verbs with vowel [ə]

Base form of verb	Past form of verb
məg 'bury'	*mɨg-ə*
ɓəm 'damage'	*ɓɨm-ə*

In both data sets in Tables 3 and 4, the past form again has the high vowel [ɨ] in the root, and the suffix consistently has the mid vowel [ə], rather than the low vowel [a] seen in Table 2. I noted earlier that it is not easy to see whether the central vowels seen on the surface in Table 3 and 4 are the same as the underlying vowels of these verbs, because of the centralizing influence of the closed syllable structure CVC in these lexemes. Before concluding anything on that score, let us continue our examination of the other vowels that occur in verbs.

Tables 5 and 6 show the behavior of the front vowels [i] and [e] in the base form, respectively:

Table 5. Base form and past form of verbs with vowel [i]

Base form of verb	Past form of verb
yig 'accept'	*yig-ə*
ʃir 'tighten'	*ʃir-ə*

DOI 10.15460/auue.2025.98.1.387

Table 6. Base form and past form of verbs with vowel [e]

Base form of verb	Past form of verb
ɲen 'lack'	*ɲin-ə*
ʃed 'tear'	*ʃid-ə*

Once again, the root vowel in the past always appears as a high vowel, here the front vowel /i/. I will claim for the time being that the underlying vowels of these verbs are /i/ and /e/, respectively, and that the initial palatal consonant protects the vowel from any centralizing influence of the closed lexeme structure CVC. As we saw in Tables 3–4, the suffix also takes the shape of the mid vowel [ə].

When we turn our attention to the back vowels, we find a parallel situation. The examples in Table 7 show the behavior of the lexical vowel /u/, and in Table 8 the behavior of /o/.

Table 7. Base form and past form of verbs with vowel [u]

Base form of verb	Past form of verb
wuj 'urinate'	*wuj-o*
gul 'touch'	*gul-o* ~ *gul-ə*

Table 8. Base form and past form of verbs with vowel [o]

Base form of verb	Past form of verb
rog 'ration'	*rug-o*
ɗol 'hear'	*ɗul-ə*

As before, the vowel of the verb in the past form is a high vowel, this time the back vowel [u] because the base vowels /u/ and /o/ are likewise back and round. There is some variation in the quality of the suffix vowel; often we find the central vowel [ə] as before, but sometimes it is [o] – at the same height as [ə], but back and round in harmony with the preceding vowel.[3]

Finally, we examine the diphthongs *ya* and *wa* which were added to the vowel chart in (3). Their behavior in the verb paradigm is very interesting, and confirms their place in the vowel system. Consider Tables 9[4] and 10:

3 In the neighboring language Mulgi, the past suffix is also subject to rounding assimilation (Maria Gustafsson, p.c.), but in that language there is no variation; the vowel is always the back round [o].

4 It should be noted that words with the diphthong [ya] (variant [ye]) are not nearly as frequent as words with the common diphthong [wa].

DOI 10.15460/auue.2025.98.1.387

Table 9. Base form and past form of verbs with vowel [ya]

Base form of verb	Past form of verb
lyag 'be equal'	*lig-a*
kyad 'scrape'	*kid-a*

Table 10. Base form and past form of verbs with vowel [wa]

Base form of verb	Past form of verb
jwar 'lose weight'	*jur-a*
dwag 'throw'	*dug-a*

In the past form, we find a simple high vowel of the same quality as the glide element in the diphthong, and the suffix vowel is once again /a/. This alternation confirms my earlier assertion that the diphthong behaves as a single vowel, since it corresponds to a simple short vowel in the past form.

To further support the treatment of the two diphthongs as units, we return to the forms for inalienable possession in the nominal system. In Tables 11 and 12 we see the same alternations as in Tables 9 and 10.

Table 11. Alternation of [ya/ye] with [i] in nouns

Noun	my …	Gloss
dyer	*dir-an*	'arm'
kya-[5]	*kir-an*	'grandmother'

Table 12. Alternation of [wa] with [u] in nouns

Noun	my …	Gloss
kwad	*kud-an*	'neck'
pwag	*pug-an*	'shoulder'

5 This example, taken from Broß (1988: 82), is evidently a bound root, and is somewhat irregular; nonetheless, the behavior of its root vowel conforms exactly to the patterns established here. The fact that affixation directly onto a noun is limited to inalienable possession (Cray 2006: 8), along with the reduced frequency of the diphthong [ya/ye], explains why there are not a lot of examples of vowel alternation to draw from in the nominal system.

DOI 10.15460/auue.2025.98.1.387

4 Operation of prosodies

How, then, are these vowel alternations observed in the verbal paradigm to be analyzed? Do we allow all nine vowels of (3) as underlying vowels for the language? And how do we best account for the very regular patterns of vowel alternations in the past forms of the verb? I begin by examining that past form.

First, I propose that the suffix vowel in the past form is indeed an underlying /ə/, the quality that we observe in the vast majority of those forms. The suffix vowel shows up as [-a] only when the vowel of the verb root is [a] or a diphthong involving [a].

Secondly, it is important to note that the vowel found in the past form of the verb root is always a high vowel, whether [ɨ] or [i] or [u]. And the specific quality of this high vowel always corresponds to the horizontal position of the vowel in the base form of the verb: [ɨ] when the base vowel is [ɨ], [ə], or [a]; [i] when the base vowel is [i], [e], or [ya]; and [u] when the base vowel is [u], [o], or [wa]. So the quality of the vowel of the stem in the past form is always predictable, and there is no longer anything distinctive about its quality. For this reason I claim that in the past form, the verb's underlying vowel has simply been deleted, and is filled in on the surface with an epenthetic high vowel of the appropriate quality.

If that be the case, then, how do we ensure that the proper front or back quality of the vowel is achieved in the surface form of the past? Consider the example of (9), repeated from Table 8. The morphological process of past formation will add the suffix |-ə| and delete the root vowel [o]. But if the root vowel disappears, there is no means of determining the backness of the epenthetic high vowel that is supplied in the surface form.

(9)	Base form	Past formation	Surface form
	ɗ o l →	ɗ Ø l - ə →	[ɗilə]? [ɗɨlə]? [ɗulə]?

This situation leads us to propose that the base vowel [o] of (9) is composed of at least two features:

(a) a feature of height (mid-ness in the case of [o]), and
(b) a feature of backness and roundness, which I will call LABIALIZATION (or simply LAB).

When the vowel of the verb is deleted, then, it is only its height feature that is deleted. In (9) the labialization feature remains, and that

DOI 10.15460/auue.2025.98.1.387

feature is what ensures that the epenthetic vowel supplied in the past form is [u], so that the correct form [ɗulə] will be obtained. In parallel fashion, front vowels such as [i], [e], and [ya] are characterized by a feature of PALATALIZATION (abbreviated PAL), in addition to the feature that specifies their height. The derivations of Table 13 show a revised conception of the process, and the quality of the epenthetic vowel in the past form is correctly specified.

Table 13. Derivations of past form of verbs containing PAL and LAB

Base form		Past formation		Surface form
s ^{PAL}e d	→	s ^{PAL}Ø d - ə	→	[ʃidə]
ɗ ^{LAB}o l	→	ɗ ^{LAB}Ø l - ə	→	[ɗulə]

I point out here that the two features PAL and LAB are the same as the prosodies that have been posited in the analysis of numerous Chadic languages (Wolff 1981, Barreteau 1987, Roberts 2001), in particular for languages of the Central branch. Roberts (2009, 2025) also shows that they also operate in a limited way in some Eastern Chadic languages. The present analysis shows that Ndam can be added to the list of languages that rely on prosodies in the operation of their vowel system. A prosody in Chadic languages is a distinctive suprasegmental feature which associates to certain morphemes as a whole. The realization of the prosody varies widely from language to language, according to language-specific rules (Roberts 2001). In some languages, the prosody spreads to most or all syllables of the word, and often starts from a specific docking point. In others, the effect of the prosody does not usually extend beyond a single syllable, as in Mawa (Roberts 2009). In any case, though, no more than one instance of the prosody is allowed per morpheme. For Ndam, I will assume that the domain of PAL and LAB is the whole word, although their principal effect is on the vowel of the first syllable.[6] In the case of LAB, which seems to be the stronger of the two prosodies, its effect optionally continues beyond the root to round the suffix vowel, as seen in the past tense forms of the verbs in Tables 7 and 8.

Now we turn again to the mid vowels [e] and [o]. I am proposing that each of these is characterized by two principal features, either LAB (in the case of [o]) or PAL (in the case of [e]), and also a feature

6 I signal one effect of the PAL prosody on consonants, nonetheless. A lexeme which contains the fricative /s/ and the pal prosody will realize /s/ as its allophone [ʃ], as in [ʃed] the first example of Table 13, or [leʃ] 'to cover', past form [liʃə].

DOI 10.15460/auue.2025.98.1.387

of mid-ness. Without entering into a detailed debate of the feature system that most appropriately characterizes these vowels, I will simply represent "midness" by the central vowel /ə/, which I will henceforth consider to be an underlying vowel of the Ndam system.

The low vowels in the system, as we have presented it, are three: [a], plus the two diphthongs [ya] and [wa]. I now claim that the two diphthongs are simply the realization of the underlying vowel /a/ when associated with the features PAL and LAB, respectively, that we have already argued for in the case of the mid vowels.

As for the three high vowels [i], [ɨ], and [u], I claim that they are not underlying vowels of the system at all, but are rather epenthetic. The consequence of this position is that certain lexemes of the language simply have no vowel at all in their underlying structure. This leads to a rather abstract conception of the vowel system and of the underlying structure of Ndam lexemes, it is true. Nonetheless, this conception allows us to take many regularities and characteristics of the language into a unified account.

A final word is necessary with respect to the suffix vowel of the past form of the verb, which evidently changes to [a] when the root vowel involves the low vowel /a/. One way to account for this vowel is to claim that the vowel [a] carries a feature [+low], which spreads from the stem to the suffix vowel. In the process of past formation, the stem vowel is deleted, but its low height spreads to the suffix. The surface vowel [a] in the past tense forms of [kɨla, gɨda] in Table 2 is simply the realization of the [+low] feature which replaces the mid height of the underlying vowel [ə], and is displayed graphically in Fig. 1.

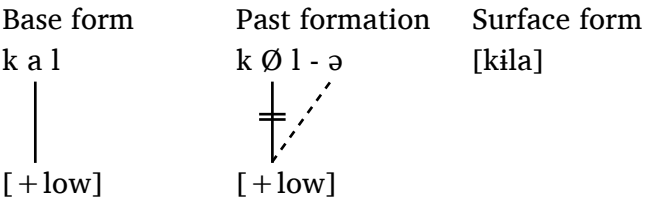


Figure 1. Specification of the suffix vowel [a]

Consider now the complete array of the verb paradigms that we have been examining. The first set of verbs, whose derivations are shown in Tables 14 and 15, is the most straightforward and involves no prosodies. The three verbs chosen have each of the three possibilities in the vowel position in the root: no vowel (Ø), /ə/, and /a/. In each

DOI 10.15460/auue.2025.98.1.387

case the abstract underlying form is shown, along with the processes involved to produce the surface form. Table 14 shows the derivation of the base form of each of the verbs; the only phonological process observable here is the epenthesis of the high vowel [ɨ] into the verb that has no underlying vowel.

Table 14. Derivation of the base form of verbs with no prosodies

	Base	Base	Base
Abstract form	g Ø d	m ə g	k a l
Epenthesis	g ɨ d	—	—
Surface form	[gɨd]	[məg]	[kal]
	‘bite’	‘sneeze’	‘see’

The derivation of the past tense forms of these same verbs is shown in Table 15. As indicated before, the past is constructed by adding the suffix *-ə*, shown in the initial abstract form, and by deleting the stem vowel. Before the stem vowel is deleted, however, verbs with the vowel /a/ cause the suffix vowel to lower. At this point, all three verbs have no vowel in the root, so again the vowel [ɨ] is epenthesized to allow them to be pronounced.

Table 15. Derivation of the past form of verbs with no prosodies

	Past form	Past form	Past form
Abstract form	g Ø d – ə	m ə g – ə	k a l – ə
Lowering	—	—	k a l – a
Deletion	—	m Ø g – ə	k Ø l – a
Epenthesis	g ɨ d – ə	m ɨ g – ə	k ɨ l – a
Surface form	[gɨdə]	[mɨgə]	[kɨla]

The verbs of the second set, shown in Table 16 with the corresponding past forms in Table 17, all have the underlying prosody of labialization, but are otherwise completely parallel to the previous set. The only complexity is the phonetic realization of the prosody as it fuses with the root vowel, creating [u] from [ɨ], [o] from underlying /ə/, and the diphthong [wa] from /a/.

DOI 10.15460/auue.2025.98.1.387

Table 16. Derivation of the base form of verbs with the LAB prosody

	Base	Base	Base
Abstract form	LAB g Ø l	LAB d ə y	LAB ɗ a g
Epenthesis	LAB g ɨ l	—	—
Surface form	[gul]	[doy]	[ɗwag]
	‘touch’	‘read’	‘throw’

Table 17. Derivation of the past form of verbs with the LAB prosody

	Past form	Past form	Past form
Abstract form	LAB g Ø l – ə	LAB d ə y – ə	LAB ɗ a g – ə
Lowering	—	—	LAB ɗ a g – a
Deletion	—	LAB d Ø y – ə	LAB ɗ Ø g – a
Epenthesis	LAB g ɨ l – ə	LAB d ɨ y – ə	LAB ɗ ɨ g – a
Surface form	[gulə]	[duyə]	[ɗuga]

The final set, displayed in Tables 18 and 19, again shows the behavior of the three root vowels /Ø, ə, a/, but this time when the prosody of palatalization is present. This prosody creates [i] from the epenthesized [ɨ], [e] from the mid vowel /ə/, and the diphthong [ya] from /a/.

Table 18. Derivation of the base form of verbs with the PAL prosody

	Base	Base	Base
Abstract form	PAL k Ø n	PAL s ə d	PAL l a g
Epenthesis	PAL k ɨ n		
Surface form	[kin]	[ʃed]	[lyag]
	‘release’	‘tear’	‘be equal’

Table 19. Derivation of the past form of verbs with the PAL prosody

	Past form	Past form	Past form
Abstract form	PAL k Ø n – ə	PAL s ə d – ə	PAL l a g – ə
Lowering	—	—	PAL l a g – a
Deletion	—	PAL s Ø d – ə	PAL l Ø g – a
Epenthesis	PAL k ɨ n – ə	PAL s ɨ d – ə	PAL l ɨ g – a
Surface form	[kinə]	[ʃidə]	[liga]

A more radical analysis of the past tense form is to claim that its suffix vowel is simply an empty V slot, and that the verb’s underlying root vowel, either /ə/ or /a/, is transferred to fill that slot before

DOI 10.15460/auue.2025.98.1.387

that vowel is deleted in its position in the stem. Then, as before, to make the word pronounceable, an epenthetic [ɨ] must be inserted, which is further subject to modification by LAB or PAL, if one of those prosodies is present. The complication with this analysis is for verbs which have no vowel to begin with; in those cases there is no vowel to transfer to the suffix V slot. These verbs have a suffix in the past tense form nonetheless, whose quality is [ə]. To follow this analysis, one would have to assume that the quality of this empty vowel is filled with [ə] as a default. This conception is graphically represented in Fig. 2, which shows the derivation of the past tense form of each of the three verbs of Table 15. The absence of a vowel, represented by Ø (where the vowel has been deleted), is filled by the epenthetic [ɨ], and the empty vowel slot, represented by V, is filled by a default [ə] if its quality is not specified in any other way.

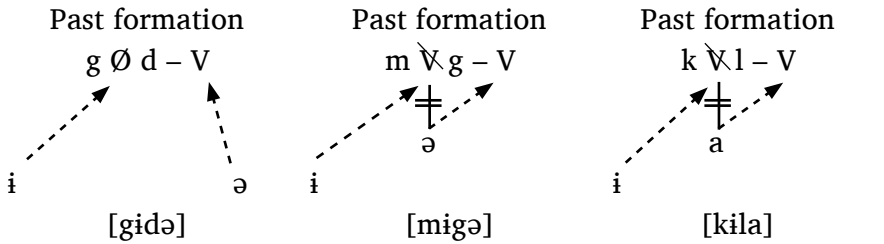


Figure 2. Derivation of the past form by transfer of the root V to the suffix

However, because of the complications involved in this latter analysis, I retain the earlier analysis, as presented in Fig. 1 and in the derivations of Tables 14–19.

5 The Ndam vowel system in a wider Chadic perspective

I conclude by summarizing my understanding of the underlying vowel system of Ndam, then considering the implications of this analysis for Chadic vowel systems in general. A final note warns of the dangers of confusing the vowels [ə] and [ɨ] in the transcription of Chadic languages.

5.1 The underlying vowels of Ndam

The result of all the preceding analytic operations leaves us with the following understanding of the underlying vowel system of Ndam, shown in (10):

DOI 10.15460/auue.2025.98.1.387

(10) LAB PAL

ə

a

There are only two underlying vowels, /ə/ and /a/. The two prosodies LAB and PAL, while not vowels, must also be recognized as distinctive underlying phonological units in the language. Together these basic units can explain the complete range of the nine surface vowels of language, originally presented above in (3). They also make sense of the complex morpho-phonological alternations that we observe in the verbal and nominal systems. I should add that although the paradigms from the possessive nouns are not presented in full in this paper, it can be shown that they follow the same patterns that we have observed in the verb data.

5.2 Vowel height in Chadic languages

The findings in Ndam confirm a generalization about the phonological levels of height that is respected by all Chadic vowel systems. It can be stated that all Chadic systems recognize only two fundamental levels, high and non-high. High vowels are either epenthetic and completely non-phonological, as we have seen in Ndam, or else they are severely limited in their distribution and in their phonological range of operation. This is particularly true of languages which have been analyzed as having /a/ as the only underlying vowel. Non-high vowels, on the other hand, have real phonological status: they are full phonemes, and their behavior is quite distinct from that of the high vowels. In languages like Ndam, we are forced to recognize two levels of non-high vowels, represented by the mid vowel /ə/ and the low vowel /a/, but the basic divide between high and non-high vowels remains. This claim of two fundamental levels of height can be substantiated in numerous Central Chadic languages, but also in a number of Eastern Chadic languages, namely Ndam, Somrai, Mulgi, Mukulu, Migaama, Mubi, and others; see also Wolff forthcoming and Roberts 2025.

I suspect that the development of a height distinction among the non-high vowels in some Eastern Chadic languages like Ndam may be a relatively recent development. There is evidence that some languages of the Guéra region like Dangaleat and Sokoro, for example, may be starting to phonologize a distinction between [e] and [ɛ] and

DOI 10.15460/auue.2025.98.1.387

between [o] and [ɔ]. Nonetheless, they still keep all of those non-high vowels quite distinct from the high vowels.

All Chadic vowel systems are rectangular in nature, as opposed to Bantu systems, for example, which are normally triangular. The display of (11) shows the basic grid for many languages which clearly need only two levels; this includes many languages of the Guéra group in Eastern Chadic, as well as those of the Masa and Central Chadic branches. The display of (12) is appropriate for those languages which have developed a secondary distinction among the non-high vowels, like Ndam. In both types, however, a strict division is maintained between the high vowels and the non-high vowels, represented by the thick bar between the two levels.

(11) Vowel system with two height levels

(i)	(ɨ)	(u)
e / ɛ	a	o / ɔ

(12) Vowel system with three height levels

(i)	(ɨ)	(u)
e	ə	o
i̯a / ɛ	a	u̯a / ɔ

5.3 The ambivalent use of <ə>

In closing, I find it necessary to comment on a long-standing tradition in Chadic linguistics, namely the use of the schwa symbol <ə> which is often seen in phonetic and phonological transcriptions. In many cases, it is clear that the schwa is intended to represent a high central vowel, which would be [ɨ] according to the IPA. In other cases it is not clear whether the intended vowel is the high [ɨ], or else a mid vowel, which is the approved use of the symbol [ə] according to the IPA. There are numerous languages in the Eastern branch of Chadic languages, like Ndam, in which both the high vowel [ɨ] and the mid vowel [ə] are found on the surface, and the two behave very differently. In these cases we cannot afford to be ambiguous in the use of symbols for transcription. It is possible that earlier researchers could not hear the difference between [ɨ] and [ə], and if that were the case, it would spell disaster for their analysis of languages like

DOI 10.15460/auue.2025.98.1.387

Ndam, where the two vowels both occur, but with radically different statuses: [ə] is a "real" phonemic vowel, whereas [ɨ] is merely epenthetic and non-phonemic. In many Central Chadic languages there is only one non-low central vowel, and one could reckon that the use of the schwa symbol to represent it is tolerable. Nonetheless, even if there is phonetic variation in the realization of the vowel transcribed with <ə>, the essential question, as I note above, is whether the vowel belongs phonologically to the high set, or to the non-high set. I know it is difficult to change long-standing habits, but henceforth I would encourage all Chadicists to adhere as closely as possible to IPA standards in phonetic transcription, in order to avoid any possible confusion.

References

Barreteau, Daniel. 1987. Du vocalisme en tchadique. In Daniel Barreteau (ed.), *Langues et cultures dans le bassin du lac Tchad.* Paris: ORSTOM. 161–191.

Boyeldieu, Pascal. 1985. *La langue lua ("niellim")*. Paris: SELAF.

Broß, Michael. 1988. *Materialien zur Sprache der Ndam von Dik (Republik Tschad)*. Marburg: Philipps-Universität Marburg Master's thesis.

Cray, William. 2006. *A grammar sketch of Ndam.* N'Djaména: AIM. Ms.

Cray, William. 2012a. *Lexique Dam-Français.* Ms.

Cray, William. 2012b. *Ndam verb inflections.* Ms.

Frajzyngier, Zygmunt. 1983. The underlying form of the verb in Proto-Chadic. In Herrmann Jungraithmayr (ed.), *The Chad languages in the Hamitosemitic Nigritic border area.* 123–143.

Jungraithmayr, Herrmann. 1977. Apophony and grammatical tone in the tense system of Chadic languages. *Afrika und Übersee* 60. 79–82.

Lionnet, Florian. forthcoming. *The phonology of Laal.* Ms.

Maddieson, Ian. 1984. *Patterns of sounds.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511753459

Newman, Paul. 2013. The Chadic language family: classification and name index. Mega-Chad Research Network: Electronic publication. https://www.cepam.cnrs.fr/megatchad/publications/Newman-2013-Chadic-Classification-and-Index.pdf

Roberts, James. 2001. Phonological features of Central Chadic languages. In Ngessimo M. Mutaka & Sammy B. Chumbow (eds.), *Research mate*

DOI 10.15460/auue.2025.98.1.387

in African linguistics: Focus on Cameroon. Köln: Rüdiger Köppe Verlag. 93–118.

Roberts, James. 2009. Palatalization and labialization in Mawa (Eastern Chadic). In Eva Rothmaler (ed.), *Topics in Chadic linguistics V: Comparative and descriptive studies.* Köln: Rüdiger Köppe Verlag. 129–140.

Roberts, James. 2025. Palatalization and labialization in the Chadic languages of Chad. *Afrika und Übersee* 98. 41–62.

Wolff, H. Ekkehard. 1981. Vocalisation patterns, prosodies, and Chadic reconstructions. In William Ronald Leben (ed.), *Précis from the Twelfth Conference on African Linguistics*. Los Angeles: Department of Linguistics, UCLA. 144–148.

Wolff, H. Ekkehard. 2024. *Lexical reconstruction in Central Chadic: a comparative study of vowels, consonants and prosodies.* Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009346375

Wolff, H. Ekkehard. forthcoming. Typology and evolution of minimal vowel systems in Central Chadic (Afroasiatic). In Natalia Kuznetsova, Cormac Anderson & Shelece Easterday (eds.), *Rarities in phonetics and phonology: evolutionary, structural, typological and social dimensions.* Berlin: Language Science Press.

DOI 10.15460/auue.2025.98.1.387

Comparing phasal aspect in some West Chadic and West Benue Congo languages: Inching toward a crosslinguistic West African perspective

Ronald P. Schaefer[a] *& Francis O. Egbokhare*[b]

Southern Illinois University Edwardsville[a] & University of Ibadan[b]
rschaef@siue.edu
francis.egbokhare@gmail.com

Abstract
We undertake an initial crosslinguistic comparison of formal coding linked to the functions of phasal aspect in selected languages from West Chadic and West Benue Congo. Canonical phasal aspect types are semantically characterized as INITIATION, TERMINATION, and CONTINUATION. We consider phasal and non-phasal forms, the latter having similar positional limitations as phasal aspect expressions. In Hausa and Miya, phasal aspect is coded by verbs and an immediately following complement. Hausa phasal verbs take truncated sentence-complements, gerundives or participles, or sentence-complements, while those in Miya show truncation or parataxis. In the same slot as phasal aspect forms Hausa exhibits verbs that do not convey aspectuality. Instead, they provide either temporal interpretations of various kinds, frequency, duration, habituality, or meanings that are capacitive or quantitative in nature. In Emai of West Benue Congo, phasal aspect is coded by verbs with truncated complements and by grammatical morphemes. The latter code canonical phasal aspect and precede the main verb. In the same pre-verb position, Emai and other West Benue Congo languages exhibit grammatical morphemes that show non-canonical phasal aspect: temporal interpretations as well as capacitive, volitive, or quantitative meanings. We conclude that canonical phasal aspect forms and similarly positioned but non-phasal forms in West Africa provide fertile ground for further crosslinguistic comparison.

Keywords: Phasal aspect, Hausa, Miya, Edoid, Yoruboid, preverbs

1 Introduction

Variation certainly characterizes the Chadic language family from east to west. However, phasal aspect forms in West Chadic exhibit rela-

DOI 10.15460/auue.2025.98.1.356

tively uniform syntax and semantics. They show forms and meanings linked most often to truncated complements, as analyzed by Newman (2000), Jaggar (2001), and Schuh (1998). Still, their findings have not been incorporated into a broader crosslinguistic assessment with non-Chadic languages of West Africa. To inch toward such comparison with West Benue Congo, we examine Hausa verbs that express phasal aspect and associated meanings. We outline their potential semantic classes, assess their grammatical complement types, and compare them to West Benue Congo Emai (Edoid) and Yoruba.

2 Aspectual verbs in Hausa of West Chadic

We begin with semantic classes. Investigations of phasal aspect by Dixon (1991: 172, 2010: 417) and Levin (1993: 274–275) have established an initial tripartite baseline of canonical semantic types. They identified three types: INITIATION, TERMINATION, and CONTINUATION. Based on Newman (2000), Hausa shows three phasal aspect verbs that are clearly canonical: *fara* 'begin/start', *gama* 'finish', and *riƙa* 'continue'. Each of these verbs in a clausal structure is illustrated in (1–3).

(1) Hausa (Newman 2000: 67)
Yā *fārà* *yî-n* *màganā.*
3SG.M.COMP begin do-GEN talk
'He began/started talking.'

(2) Hausa (Newman 2000: 69)
Màtā-tā *gamà* *dafà* *àbinci-n* *rāna.*
wife.GEN-1SG.F-COMP finish cooking food-GEN midday
'My wife finished cooking lunch.'

(3) Hausa (Jaggar 2001: 549)
Zā *mù* *riƙà* *kỏyo-n* *Hausa* *à* *jāmi'àř.*
FUT 1PL continue learn-GEN Hausa LOC university
'We will continue to study Hausa at the university.'

Overall, forty-one phasal aspect expressions are identified for Hausa by Newman (2000), thirty-five by Jaggar (2001), and in Miya fourteen are found in Schuh (1998). Several of these forms are not as

DOI 10.15460/auue.2025.98.1.356

strongly delimited in their meaning as the tripartite division of Dixon (1991) and Levin (1993) suggests. For instance in Hausa, there is *koma* 'resume' and *fasa* 'postpone'. They characterize event INITIATION but only in terms of temporal resumption and delay, respectively. Both assume the existence and happening of an event. The verb *kara* 'repeat/do again' bears on CONTINUATION, albeit of a discontinuous sort, since it refers to repetition of an event. The verb *samu* 'manage to/succeed in doing' and its negative opposite *rasa* 'not manage' augment dimensions of TERMINATION, not simply confirm it. The verb forms *kusa* 'almost, be about to' and *nèmā* 'try to, be about to' assume INITIATION as background for aspectually proximate meanings.

Further examination reveals verbs that hold the same position as Hausa phasal aspect verbs but do not convey either of the three canonical types, even when the latter are construed broadly. Forms in this second tier, the non-phasal forms, do not highlight an event edge (INITIATION/TERMINATION) or non-edge CONTINUATION. Second tier verbs characterize an event's temporal nature regarding frequency, duration or habituality. Among non-phasal forms that characterize event frequency, there are verbal expressions in Hausa such as *cika* 'do too much', *faye* 'do too often', and *rage* 'do less often'. Others signal the habituality of an event in the present or past: respectively, *fi* 'do something usually' vs. *sābấ (dà) / saba da* 'used to do'. Still other verbs characterize event duration that remains indefinite. Included are *jima da* 'spend some time / a while', *dade* 'spend long time', and *rabu da* 'time has passed since'.

A third tier of non-phasal aspect forms in the West Chadic languages under review are oriented toward grammatical subjects. Relevant forms reveal meanings in Hausa that are capacitive (*isa* 'be capable of', *iya* 'be able to') or quantitative (*tara* 'share doing, do together'). In addition to the broad semantic range of expressions included under phasal aspect in Hausa, there are collocations that evince a non-literal, conventional metaphoric character ('eat at the front'), as in (4). We can conclude that there are verbs in Hausa that are syntactically positioned like phasal forms but are not confined to the canonical phasal meanings outlined by Levin (1993) and Dixon (1991).

DOI 10.15460/auue.2025.98.1.356

(4) Hausa (Newman 2000: 65)
Shāgāri yā ci gāba dà zamā
Shagari 3SG.M.COMP eat front COMP being
shugàba-nmù.
president-GEN.1PL
'Shagari continued to be our president.'

We now briefly consider phasal aspect complements in the West Chadic languages under review. Overall, complement form and grammatical marking manifest some variation. Complements are often nonfinite and deverbal, with dominant examples being infinitives and gerundives, as in the examples of (1–3), rather than finite clauses. In Hausa, complements are marked by comitative *dà*; it has the sense 'with' when combined with deverbals but the sense 'that' when combined with sentence complements.

Complements of phasal aspect verbs in Hausa partially contrast with those in Miya. Schuh (1998) notes the absence of sentence- complement markers and the prevalence of simple juxtaposition (parataxis) or of adpositional phrases (prepositional) showing *à̀* 'with'. This absence characterizes Miya phasal and non-phasal aspect verbs (5–7).

(5) Miya (Schuh 1998: 354)
míy fārà táwàzə bə́nà.
1PL begin planting yesterday
'We began planting yesterday.'

(6) Miya (Schuh 1998: 355)
míy sə́ mbə́ ghànakə súwày.
1PL IPFV finish building tomorrow
'We will finish building tomorrow.'

(7) Miya (Schuh 1998: 354)
mən ā sə́kɛ̀ bəsaka nakən ndúwul.
1SG IPFV repeat washing this pot
'I will wash this pot again.'

We can sum up as follows. The Chadic languages under review display verbs that code the canonical division of phasal aspect described

DOI 10.15460/auue.2025.98.1.356

by Dixon (1991) and Levin (1993). In addition to phasality, Chadic verbs, positionally related to the canonical phasal aspect forms, code meanings that are neither edge-inclusive (INITIATION, TERMINATION) nor edge-exclusive (CONTINUATION). Instead, these additional verbs code event temporal frequency, temporal qualifications that are indefinite, conditions of negation as well as capacitive and quantitative attributes of a grammatical subject. And still other expressions of phasal and non-phasal aspect utilize conventional metaphor.

3 Phasal aspect in Emai of West Benue Congo

A crosslinguistic perspective may shed further light on the structural patterns gleaned from West Chadic. Among West Benue Congo languages, phasal aspect verbs can be identified. However, Edoid, Yoruboid, and Oko (where orthographic o = ɔ) of the Nupoid group code the phasal aspect domain in a slightly different fashion from West Chadic. They do so in Edoid Emai with grammatical morphemes, not simply verbs (Schaefer & Egbokhare 2007, 2017). For example, Emai verbs do not code the three canonical semantic classes of phasal aspect advanced by Levin (1993) and Dixon (1991). It is grammatical morphemes preceding the main verb that code each canonical phasal aspect type.

3.1 Phasal aspect verbs in Emai

Emai shows verbs with phasal aspect meanings that are intransitive, transitive, or labile. Among exclusively intransitive verbs is *bɛɛ* ‘start’, which refers to INITIATION. As its argument, *bɛɛ* takes a gerundive complement marked by adposition *vbi* (8a) or a sentence complement marked by indicative complementizer *khi* (8b–c). The latter imposes a same-subject relation on the complement clause relative to the matrix clause.

(8) Emai (Schaefer & Egbokhare 2007: 62)

a. *ɔ́lí*	*ɔ́mɔ̀hè*	*bɛ́ɛ́*	*vbí*	*émáé*	*úèmí.*
ART	man:PRX	PST:start:PFV	LOC	food	eating

‘The man started eating food.’

DOI 10.15460/auue.2025.98.1.356

b. *ɔ́lí ɔ́mɔ̀ bɛ́ɛ́ khì ɔ̀ ɔ́*
ART child:PRX PST:start:PFV IND SM PRS.PROG
khùòkhúó.
crawling
'The baby started crawling.'

c. *ɔ́lì èràìn bɛ́ɛ́ khì ɔ̀ ɔ́ tó.*
ART fire:PRX PST:start:PFV IND SM PRS.PROG burn
'The fire started burning.'

Another intransitive-only verb is *hɛna* 'cease, abstain, stop'; it bears on TERMINATION. The verb *hɛna* has no following argument. It occurs with a following locative marked gerundive. But *hɛna* retains the perfective suffix *-í* and its metatonic properties (Hyman & Lionnet 2012, Schaefer & Egbokhare 2021). Under this condition, the locative phrase serves as an adjunct, rather than argument (9a–b).

(9) Emai (Schaefer & Egbokhare 2007: 167)

a. *ókhɔ̀ìn hɛ́ná-ì.*
fighting:PRX PST:cease-PFV
'The fighting has ceased / stopped.'

b. *òjè hɛ́ná-í vbí ényɔ́ údàmí.*
Oje:PRX PST:cease-PFV LOC wine drinking
'Oje ceased / stopped / abstained from drinking wine.'

Emai also exhibits labile verbs that convey TERMINATION. They reject complements that are marked by either adpositions or sentence-complementizers. Instead, there is a positional tendency for an activity or abstract nominal as subject of the intransitive form or direct object of the transitive form (10a–b, 11a–b, 12a–b, 13a–b).

(10) Emai (Schaefer & Egbokhare 2007: 132)

a. *ɔ́lì òbìà fóó-ì.*
ART work:PRX PST:finish-PFV
'The work has gotten finished.'

DOI 10.15460/auue.2025.98.1.356

b. *òjè fóó ɔ́lì òbìà.*
Oje:PRX PST:finish:PFV the work
'Oje has finished the work.'

(11) Emai (Schaefer & Egbokhare 2007: 132)
a. *ébóó òjè fóó-ì.*
relations Oje:PRX PST:eliminate-PFV
'Oje's relations have perished / gotten eliminated.'

b. *è fóó ébóó òjè.*
3PL:PRX PST:eliminate:PFV relations Oje
'They have eliminated / finished off Oje's relations.'

(12) Emai (Schaefer & Egbokhare 2007: 132)
a. *ìbɔ̀bɔ̀dí ísì òjè fóó-ì.*
cassava ASS Oje:PRX PST:eliminate:PFV
'Oje's cassava has gotten eliminated/destroyed.'

b. *ɔ́lí ívàn fóó ójé ìbɔ̀bɔ̀dí.*
ART grasscutter:PRX PST:eliminate:PFV Oje cassava
'The grasscutter eliminated/destroyed Oje's cassava.'

(13) Emai (Schaefer & Egbokhare 2007: 243)
a. *ɔ́lì òsí↓é kúyé ↓á.*
ART performance:DST PST:eliminate:PFV ITV
'The performance has become terminated.'

b. *òjè kúyé ɔ́lì òsí↓é á.*
Oje:PRX PST:eliminate:PFV ART performance ITV
'Oje has terminated / broke up the performance.'

Regarding the overall phasal aspect domain, Emai shows verbs that are polysemous. Some are exclusively intransitive, others are not. One example of such a verb pertains to the canonical phasal aspect of TERMINATION. For instance, the intransitive verb *muzan* has the phasal aspect meaning 'halt, stop' (14a–b) as well as the posture sense 'stand' (14c).

DOI 10.15460/auue.2025.98.1.356

(14) Emai (Schaefer & Egbokhare 2007: 282)

a. *òjè* *múzán-ì.*
Oje:PRX PST:halt-PFV
'Oje has halted / stopped.'

b. *ɔ́lì* *òkpòsò* *nwú* *ɔ́lì* *ìmátò* *múzán.*
ART woman:PRX PST:take:PFV ART car halt
'The woman has made the car stop / hailed the car.'

c. *òjè* *múzán-í* *vbí* *úkpódɛ̀.*
Oje:PRX PST:stand-PFV LOC road
'Oje has stood on the road.'

Verbs of causative motion in Emai exhibit exclusively transitive forms. They include *khu* 'chase', *roo* 'release', and *gbaan* 'wind', which have phasal aspect meanings pertaining to TERMINATION ('stop' for *khu* in 15, 'cease' for *roo* in 16) or INITIATION ('start' for *gbaan'* in 17).

(15) Emai (Schaefer & Egbokhare 2007: 248)

a. *èɡè* *khú* *ójé* *vbí* *ɛ́nyɔ́* *údàmí.*
Ege:PRX PST:chase:PFV Oje LOC wine drinking
'Ege stopped Oje from drinking wine.'

b. *èɡè* *khú* *òjè.*
Ege:PRX PST:chase:PFV Oje
'Ege has chased Oje.'

(16) Emai (Schaefer & Egbokhare 2007: 357)

a. *òjè* *róó* *óbɔ́* *vbí* *ɛ́nyɔ́* *údàmí.*
Oje:PRX PST:release:PFV hand LOC wine drinking
'Oje ceased drinking.' (Lit. 'Oje released his hand from wine drinking.')

b. *òjè* *róó* *ɛ́ɡé* *óbɔ̀.*
Oje:PRX PST:release:PFV Ege arm
'Oje released Ege.' / 'Oje left Ege alone.'

Each TERMINATION form in (15–16) allows adposition complements that contain a gerundive phrase. However, verb *gbaan*, which

DOI 10.15460/auue.2025.98.1.356

otherwise has the meaning ‘wind’ (17b), and ventive preposition re collocate to express the phasal aspect meaning ‘start’, reflecting INITIATION. They take a gerundive in direct object position (17a).

(17) Emai (Schaefer & Egbokhare 2007: 145)

a. *yán* *ꜜgbáán* *ɔ́ká* *úkɔ̀mí* *é.*
3PL:DST PST:wind:PFV maize planting VEN
‘They started maize planting.’ (Lit. ‘They wound up maize planting.’)

b. *ɔ̀jè* *ɔ̀* *ɔ́* *gbààn* *òú.*
Oje:PRX SM:PRX PRS:PROG wind:IPFV thread
‘Oje is winding thread.’

Still other verbs that engage meanings related to canonical phasal aspect utilize a more complex syntax or exhibit a meaning related to the negation of CONTINUATION. The verb series complex *daa nyɛ* ‘detain’ with no gerundive or sentence complement shows an obligatory extentive structure, where extentive refers to a change of position that is determined to be maximal (18a–b).

(18) Emai (Schaefer & Egbokhare 2007: 71)

a. *ìsɔ́jà* *dáá* *ɔ́ì* *nyɛ́* *vbí*
soldiers:PRX PST:reposition:PFV 3SG constrain LOC
úkpódɛ̀.
road
‘Soldiers detained him on the road / stopped him on the road.’

b. *ɔ̀jè* *dáá* *álékɛ́* *ɔ́bɔ̀* *nyɛ́.*
Oje:PRX PST:reposition:PFV Aleke arm constrain
‘Oje detained Aleke.’ / ‘Oje prevented Aleke from leaving.’

The Emai verb *hɛɛ* in construction with itive form a has only an intransitive shape. Depending on its subject collocate it has the sense ‘dissipate’ or ‘evaporate’, both of which relate to negation of a CONTINUATION condition (19a–b).

DOI 10.15460/auue.2025.98.1.356

(19) Emai (Schaefer & Egbokhare 2007: 166)

a. *ókhɔ̀ìn hɛ́ɛ́ ↓á.*
fighting:PRX PST:dissipate:PFV ITV
'The fighting has stopped. / The fighting has dissipated.'

b. *ɔ́lì àmɛ̀ hɛ́ɛ́ ↓á.*
ART water:PRX PST:dissipate:PFV ITV
'The water has evaporated. / The water has oozed away.'

Last among these verbs is *waa* 'refrain from' (as a taboo). It evinces only a transitive form. It does not exhibit an edge-inclusive sense or negation of an edge-inclusive sense. Instead, its meaning pertains to negation of an assumed durative condition, i.e. negation of CONTINUATION, where for example okra or some other food source, for instance, would normally be consumed and not be barred from consumption as a taboo item (20).

(20) Emai (Schaefer & Egbokhare 2007: 434)
ójé ɔ́ ɔ̀ wàà íshàvbɔ́.
Oje:DST SM:DST PRS.HAB refrain:IPFV okra
'Oje refrains from okra.'

3.2 Phasal aspect grammemes in Emai

In addition to verbs in the preceding section, Emai articulates meanings that reflect canonical phasal aspect with grammatical forms or grammemes. Relevant exponents are preverbal grammatical morphemes that express CONTINUATION ('continue to') or, in indicative mood, proximate relations of INITIATION ('nearly started') or TERMINATION ('nearly finished').

Of these, CONTINUATION is coded by the durative preverb form *sɛ* 'continue' (21a–b).

(21) Emai (Schaefer & Egbokhare 2017: 118)

a. *ɔ́lí ɔ́mɔ̀hè sɛ́ɛ̀ é vbí ɔ́lí émàè.*
ART man:PRX PST:DUR eat:PFV LOC ART food
'The man has continued to eat from the food.'

b. *ɔ́lí ɔ́mɔ̀hè ɔ̀ ɔ́ sɛ́ kp*
ART man:PRX SM:PRX PRS.PROG DUR wash:IPFV

DOI 10.15460/auue.2025.98.1.356

ɔ́lì ìtásà.
ART plate
'The man is continuing to wash the plate.'

Coding INITIATION and TERMINATION, respectively, are egressive *yà* 'nearly started' (22a–b) and ingressive *mɔ* 'nearly finished' (23a–b).

(22) Emai (Schaefer & Egbokhare 2017: 122)

a. *ɔ́lí ɔ́mɔ̀hè yà é ɔ́lí émàè.*
ART man:PRX PST:IG eat:PFV ART food
'The man nearly started eating the food (but didn't).'

b. *ɔ́lí ɔ́mɔ̀hè yà gbé ɔ́lí ákhè á.*
ART man:PRX PST:IG break:PFV ART pot itive
'The man nearly started to break the pot (but didn't).'

(23) Emai (Schaefer & Egbokhare 2017: 123)

a. *ɔ́lí ɔ́mɔ̀hè mɔ́ɔ̀ é ɔ́lí émàè lé.*
ART man:PRX PST:EG eat:PFV ART food already
'The man nearly finished eating the food already.'

b. *ɔ́lì ùbèlè mɔ́ɔ̀ vɔ̀ɔ̀n.*
ART gourd:PRX PST:EG be.full:PFV
'The gourd is nearly full.'

Phasal aspect grammemes in Emai appear formally related to synchronic verb forms, all of which are intransitive. We assume the grammemes result from the grammaticalization of verbs that once expressed motion or change of state. Preverb *sɛ* 'continue on' is related in form to path-of-motion verb *sɛ* with the sense 'move as far as / up to, reach'. It requires a locative marked oblique phrase as object (24). Corresponding to ingressive preverb *yà* 'nearly start' is the verb *ya* with the sense 'commence, start' (25). And egressive preverb *mɔ* 'nearly completed' is likely derived from a verb like *mɔ* 'bear/produce fruit', which conveys a two-stage change of state (26). These are provisional grammaticalization paths, but they strike us as promising.

DOI 10.15460/auue.2025.98.1.356

(24) Emai (Schaefer & Egbokhare 2007: 363)
òjè sɛ́ vbí ɛ́dà.
Oje:PRX PST:move.toward LOC river
'Oje has moved as far as the river.'

(25) Emai (Schaefer & Egbokhare 2007: 438)
óshàn yá-ì.
journey:PRX PST:commence-PFV
'The journey has commenced.'

(26) Emai (Schaefer & Egbokhare 2007: 280)
ɔ́lí údúkpù mɔ́-ì.
ART coconut.tree:PRX PST:bear.fruit-PFV
'The coconut tree has borne fruit.'

The distributional class of grammatical forms that express canonical phasal aspect meanings in Emai includes two forms with an additive character relative to an entire event, thus articulating neither INITIATION, CONTINUATION, nor TERMINATION. The two forms are *che* 'again' (27) and *gbo* 'also, too' (28).

(27) Emai (Schaefer & Egbokhare 2017: 118)
ɔ́lí ɔ́mɔ́hé ↓ché é émà.
ART man:DST PST:REP eat:PFV yam
'The man ate yam again.'

(28) Emai (Schaefer & Egbokhare 2017: 118)
ɔ́lí ɔ́mɔ́hé ↓gbó é ɔ́lí émàè.
ART man:DST PST:ADD eat:PFV ART food
'The man also ate the food / ate the food too.'

4 Comparing West Benue Congo and West Chadic

We turn now to a brief comparison of the coding of phasal aspect and non-phasal aspect in the West Benue Congo language Emai and the West Chadic language Hausa.

DOI 10.15460/auue.2025.98.1.356

4.1 Comparison of Edoid Emai and West Chadic Hausa

In this subsection we compare the coding of phasal aspect in Emai of Edoid and Hausa of West Chadic. An obvious difference to note from this preliminary analysis is that Emai codes phasal aspects with verb forms and grammemes, while Hausa relies primarily on verbs (although a reviewer reminds us that Hausa has at least one particle *ta* that follows optional pro-verb *yi* 'do' in *Yau an ta ruwa* 'Today it keeps on raining').

Relative to phasal aspect meanings and their coding, Emai verbs evince only partial coding. There is no verb in Emai that codes the phasal aspect notion of CONTINUATION. A second difference in coding patterns pertains to Emai usage of grammatical forms. With respect to these two devices there is a third difference. Emai conveys INITIATION and TERMINATION meanings with proximate forms that in the indicative translate with 'nearly/almost'. Table 1 summarizes the most obvious comparative relations between Hausa and Emai with respect to coding canonical phasal aspect by verbs (V) and/or grammemes (G).

Table 1. Summary of coding patterns for the three canonical meanings of phasal aspect identified by Dixon (1991) and Levin (1993)

Hausa-V		Emai-V	Emai-G
INITIATION		INITIATION	INITIATION (nearly)
TERMINATION		TERMINATION	TERMINATION (nearly)
CONTINUATION		---	CONTINUATION

Although verbs in Hausa code the canonical phasal aspect notions INITIATION, TERMINATION and CONTINUATION, non-phasal aspect is also coded by Hausa verbs. Non-phasal aspect verbs in Hausa provide temporal interpretations of event frequency, event duration, or corresponding qualifications of temporal negation. As well some non-phasal aspect verbs in Hausa code quantitative or capacitive attributes of grammatical subjects.

4.2 Comparison of Yoruba and West Chadic Hausa

In this subsection we compare the coding of non-phasal aspect in West Benue Congo Yoruba and in West Chadic Hausa. Grammatical forms of a sort that express non-phasal aspect meanings are found in pre-verb position in WBC. They occur in Edoid, Yoruboid, and Ọkọid.

DOI 10.15460/auue.2025.98.1.356

In Yoruba, non-phasal aspect forms are identified as grammatical preverbs (Bamgbose 1966, 1967; Rowlands 1969; Awobuluyi 1978). In large measure grammatical forms of this kind reflect verb qualifications discussed in a series of papers by Nuyts (2005, 2016). Preverb forms ascribe different types of properties to the grammatical subject. They express quantitative qualifications of an event such as 'together', and the mutually exclusive subcategories of dynamic modality, e.g. capacitive 'can', volitive 'mistakenly', capacitive 'courageously', temporal 'earlier on', in addition to boulomaic modality, e.g. 'without reason'.

Appearing with some equivalency to these WBC preverb meanings are West Chadic verb forms from Hausa with non-phasal aspect meanings. Table 2 provides an initial crosslinguistic identification of these non-phasal meanings and their related forms, as found in Newman (2000) for Hausa and Awobuluyi (1978) for Yoruba. Aligned with a selection of Hausa verbs are semantically similar grammatical preverbs from Yoruba of WBC. Quite similar alignments can be formulated for WBC Emai and Ọkọ.

Table 2. Non-phasal aspect forms and their meanings in Hausa and Yoruba

Category	Hausa	Yoruba
TEMP-FREQUENCY	*cikà* 'do too much of' *fayè/fiyà* 'do too much' *rage* 'do often'	
TEMP-NEGATION	*fāsà* 'postpone, fail' *gazà* 'lack' *ragè* 'do less often'	*tètè* 'without delay' *wulè* 'in vain, to no avail'
TEMP-DURATIVE	*ràbudà* 'time passed since' *jimà dà* 'be for a while' *tabà* 'ever do'	*jàjà* 'at last, finally' *kɔ́kɔ́/kɔ́* 'first'
CAPACITIVE	*iyà* 'be able to' *ìsa* 'be capable of' *sāmù* 'manage to succeed'	*lè* 'be able'
NEG CAPACITIVE	*kāsà* 'be unable to do' *rasà* 'be unable to'	
VOLITIVE		*dìídì* 'intentionally' *mɔ̀ɔ́mɔ̀* 'intentionally'

DOI 10.15460/auue.2025.98.1.356

Category	Hausa	Yoruba
BOULOMAIC		*sáà* 'without purpose' *dédé* 'without reason'
QUANTI-TATIVE	*tārà̀* 'share doing' *shā* 'do often, much'	*dìjɔ/jɔ* 'together' *jùmɔ̀* 'together'

5 Discussion

As the preceding has suggested, it is not necessarily only West Chadic forms from Hausa that convey meanings for canonical phasal aspect that may be ripe for crosslinguistic study. Non-phasal aspect forms that have been lumped together with phasal forms might also benefit from crosslinguistic study. West Chadic, more specifically Hausa, reveals event qualifications not limited to phasal notions as they are canonically construed. These include non-phasal temporal qualifications, both affirmative and negative, as well as qualifications that are capacitive or quantitative in nature. For purposes of simple comparison, we provide pre-verbal templates for Yoruba of West Benue Congo and Hausa of West Chadic.

Yoruba: PHASAL – BOULOMAIC – DYNAMIC – QUANTITATIVE

Hausa: PHASAL – TEMPORAL – CAPACITIVE – QUANTITATIVE

These templates may more sharply delineate the semantic types involved and their crosslinguistic standing relative to other event qualifications.

Obviously, we have lumped together and split up various verb meanings, many far-removed from our limited understanding of West Chadic. Whether this might be useful for crosslinguistic studies of form-function relations, and their linear order remains to be determined. But if nothing else, this exercise has called attention to areas of semantic and syntactic structure that are not uniformly aligned across languages in West Africa and so appear worthy of further crosslinguistic study.

Acknowledgements: The authors express their appreciation for the insightful and useful comments of two reviewers who willingly shared their knowledge of Hausa and the Chadic family of languages.

DOI 10.15460/auue.2025.98.1.356

Abbreviations: Abbreviations for grammatical morphemes are generally in line with those of the Leipzig Glossing Conventions. Abbreviations used in this paper are: ADD additive, ART definite article, ASS associative, COM comitative, COMP completive, DST distal temporal distance, DUR durative, EG egressive, F feminine, FUT future tense, GEN genitive, IG ingressive, IND indicative, IPFV imperfective aspect, ITV itive, LOC locative, M masculine, PFV perfective aspect, PL plural, PROG progressive, PRS present tense, PRX proximal temporal distance, PST past tense, REP repetitive, SG singular, SM subject marker, VEN ventive.

References

Atoyebi, Joseph. D. 2010. *A reference grammar of Oko: A West Benue Congo language of North-Central Nigeria.* Köln: Rüdiger Köppe Verlag.

Awobuluyi, Oladele. 1979. *Essentials of Yoruba grammar.* Ibadan: University Press Limited.

Bamgboṣe, Ayo. 1966. *A grammar of Yoruba.* Cambridge: Cambridge University Press.

Bamgboṣe, Ayo. 1967. *A short grammar of Yoruba.* Heinemann Books: Ibadan.

Dixon, Robert M. W. 1991. *A new approach to English grammar, on semantic principles.* Oxford: Clarendon Press.

Dixon. Robert M. W. 2006. Serial verb constructions: Conspectus and coda. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology.* New York: Oxford University Press. 338–350.

Dixon, Robert M. W. 2010. *Basic linguistic theory volume 2: Grammatical topics.* New York: Oxford University Press.

Elugbe, Ben. 1989. *Comparative Edoid: Phonology and lexicon.* (Port Harcourt Series 6). Port Harcourt: University of Port Harcourt Press.

Hyman, Larry M. & Florian Lionnet. 2012. Metatony in Abo (Bankon), A42. In Michael R. Marlo, Nikki B. Adams, Christopher R. Green, Michelle Morrison & Tristan M. Purvis (eds.), *Proceedings of 42nd Annual Conference of African Linguistics: African languages in context.* Somerville: Cascadilla Proceedings Project 1–14.

Jaggar, Philip J. 2001. *Hausa*. Amsterdam: John Benjamins.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation.* Chicago: University of Chicago Press.

Newman, Paul. 2000. *The Hausa language: An encyclopedic reference grammar*. New Haven: Yale University Press.

DOI 10.15460/auue.2025.98.1.356

Nuyts, Jan. 2005. The modal confusion: On terminology and the concepts behind them. In Alex Klinge & Henrik Høeg Müller (eds.), *Modality: Studies in form and function.* London: Equinox. 5–38.

Nuyts, Jan. 2016. Analyses of modal meanings. In Jan Nuyts & Johan van der Auwera (eds.), *The Oxford handbook of modality and mood*. New York: Oxford University Press. 31–49.

Rowlands, E. C. 1969. *Yoruba*. London: Hodder and Stoughton.

Schaefer, Ronald P. & Francis O. Egbokhare. 2007. *A dictionary of Emai: An Edoid language of Nigeria*. Köln: Rüdiger Köppe Verlag.

Schaefer, Ronald P. & Francis O. Egbokhare. 2017. *A grammar of Emai.* (Mouton Grammar Library Series). Berlin: De Gruyter.

Schaefer, Ronald P. & Francis O. Egbokhare. 2021. Metatony, S-constituent linkage, and cognate objects. *Proceedings of the Linguistic Society of America* 6(1). 641–648.

Schuh, Russell G. 1998. *A grammar of Miya*. Berkeley: University of California Press.

DOI 10.15460/auue.2025.98.1.356

Les relations lexicales entre le giziga – tchadique central – et le fulfulde (Cameroun)

Henry Tourneux[a] *& Antonio Michielan*[b]

LLACAN Villejuif[a]
Pontificio Istituto Missioni Estere (PIME) Milan[b]
henry.tourneux@wanadoo.fr
anto_michi3@hotmail.com

Résumé
Dans la région Extrême-Nord du Cameroun, les Guiziga sont sans doute la population la plus intimement imbriquée avec les Peuls. Ils occupent ensemble la vaste plaine du Diamaré, jusqu'aux piémonts des monts Mandara. Si leur longue cohabitation n'a pas toujours été idyllique, leurs langues respectives en ont tiré profit en s'empruntant mutuellement de nombreux vocables. En outre, le fulfulde, langue de nomades à l'origine, riche de nombreux contacts historiques, a permis à la langue giziga de recevoir des apports lexicaux de nombreuses autres langues (arabe, kanuri, hausa, anglais, etc.).

Abstract
In the Far North region of Cameroon, the Giziga are undoubtedly the population most closely intertwined with the Fulani. Together they occupy the vast Diamaré plain, as far as the foothills of the Mandara Mountains. While their long cohabitation has not always been idyllic, their respective languages have benefited by borrowing many words from each other. In addition, Fulfulde, the language of a nomadic people originally, rich in many historical contacts, has enabled the Giziga language to receive lexical contributions from many other languages (Arabic, Kanuri, Hausa, English, etc.).

Keywords: Giziga, Central Chadic, Fulfulde, borrowings, Cameroon, language contact

1 Introduction

La langue giziga appartient à la branche centrale de la famille tchadique au sein du phylum afroasiatique. Gravina (2014: 42) situe cette langue dans une sous-branche Nord, dans le grand groupe Margi-Mandara-Mofu. Le giziga appartiendrait au groupe Maroua, dans

DOI 10.15460/auue.2025.98.1.332

lequel Gravina situe deux dialectes giziga et une langue quasiment disparue (le mbazla). Les anthropologues comme Pontié (1973) distinguent effectivement deux groupes de population guiziga[1], l'un, au Nord, centré autour de Dogba, l'autre, au Sud, autour de Moutouroua. Lukas (1970), Barreteau & Dieu (2000) et Shay (2021) considèrent que les parlers de ces deux groupes, Nord et Sud, forment une seule et même langue. Barreteau & Dieu (2000: 65) leur attribuent même un coefficient de similarité de 94 %. Si les linguistes, à la suite de Lukas, ont eu tendance à accorder une place centrale au parler de Dogba, c'est probablement parce qu'ils n'arrivaient pas à imaginer l'importance passée de Maroua [Marva] pour les Guiziga, importance qui y a été occultée par la surimposition de la langue peule.

Historiquement, le groupe humain guiziga est formé de clans provenant majoritairement du Nord-Est, dont l'expansion s'est arrêtée au pied des monts Mandara (Pontié 1973 ; Guitard 2017). En tant que groupe, il semble difficile de le trouver avant le début du 18e siècle[2].

Son aire de dispersion actuelle principale est la plaine du Diamaré, qu'il se partage avec les Peuls et des établissements kanuri, tupuri, munjuk et autres.

Les Peuls sont présents dans la région à titre de pasteurs nomades depuis le 17e siècle, probablement. Mohammadou Eldridge (1976: 16), fait même remonter leur présence au Diamaré au 16e siècle. A la charnière des 18e et 19e siècles, ils ont commencé à s'organiser politiquement et, sur la lancée du jihâd d'Ousmane dan Fodio, ils ont vassalisé les populations non peules du Diamaré (Seignobos 2000b). C'est à cette époque qu'a dû commencer l'influence réciproque du fulfulde et du giziga dans le domaine lexical. Les relations sociales entre les Guiziga et les Peuls ont été, alors, tout sauf idylliques (Pontié 1981: 252–254).

Dans les pages qui suivent, nous nous basons uniquement sur les données lexicales de Michielan et al. *(pro manuscripto)*, en cours d'élaboration, fondées, au départ, sur un fichier original établi par René Jaouen. René Jaouen (1935–2002), de la congrégation des

1 Nous écrivons l'ethnonyme « Guiziga, guiziga » suivant la pratique locale courante ; en revanche, le nom de la langue « giziga » suit les règles de transcription de la phonétique internationale.

2 Communication personnelle de Christian Seignobos, par un courriel du 21 août 2021. C'est à lui aussi que nous devons la remarque sur la non-centralité de Dogba.

DOI 10.15460/auue.2025.98.1.332

Oblats de Marie Immaculée (OMI), est arrivé chez les Guiziga en 1962 et a consacré plus de trente ans de sa vie à leur étude. Nous lui devons plusieurs publications sur leur culture et sur leur langue (notamment 1990, 1995). Si le manuscrit dont nous disposons actuellement compte plus de 8 000 entrées, notre étude ne sera pourtant pas exhaustive – le domaine religieux de l'islam n'y étant pas représenté –, mais elle permettra de donner une bonne idée de la proportion des apports lexicaux respectifs dans le riche vocabulaire fourni.

Cette étude pourra ainsi contribuer à l'histoire des relations entre les populations du Diamaré et celles de ses marges occidentales. En l'absence de sources écrites suffisantes, l'étymologie peut en effet nous fournir des pistes précises (Jaggar 2010: 48) :

> « Although by no means an infallible technique, interpretation of language evidence (e.g. the etymological source of words) can provide reliable clues to the reconstruction of the culture of a people, the history of their language, and interaction with other groups and languages. »

2 Les systèmes phonologiques du giziga et du fulfulde

La phonologie du giziga est bien typique de celle des langues de la branche centrale de la famille tchadique (Wolff 2022) et le fulfulde, de celle de la famille atlantique du phylum Niger-Congo (Pozdniakov 2022).

2.1 Les consonnes du giziga

Selon E. Shay (2021: 17–28), le giziga possède 32 consonnes. Nous en dressons, d'après l'inventaire fourni par cette autrice (2021: 18), le tableau suivant dont nous modifions légèrement la structuration.

DOI 10.15460/auue.2025.98.1.332

Tableau 1. Les consonnes du giziga

	labiales	apicales	palatales	vélaires	labio-vélaires	postérieures
occlusives sourdes	*p*	*t*	*c*	*k*	*k*^{w}	
occlusives sonores	*b*	*d*	*j*	*g*	*g*^{w}	
glottalisées	*ɓ*	*ɗ*				*ʔ*
prénasalisées	*mb*	*nd*	*nj*	*ng*	*ng*^{w}	
nasales	*m*	*n*		*ŋ*		
fricatives sourdes	*f*	*s*				*h*
fricatives sonores	*v*	*z*				
latérale fricative sourde		*ɬ*				
latérale fricative sonore		*ɮ*				
continues	*w*	*l*	*y*			
battue		*r*				

2.2 Les voyelles du giziga

Voici le système vocalique, tel qu'il est présenté par Shay (2021: 25) ; le giziga comporte six voyelles phonétiques (mais seulement trois phonèmes vocaliques) :

Tableau 2. Les voyelles du giziga

i *(ə)* *u*

(e) *(o)*

a

Les voyelles entre parenthèses sont analysées comme non phonémiques. Toutes, à part la voyelle centrale haute, peuvent apparaître sous une forme phonétique longue.

DOI 10.15460/auue.2025.98.1.332

2.3 Les consonnes du fulfulde

Le fulfulde du Diamaré présente le tableau phonologique suivant pour les consonnes (Tourneux et al. 2021: 12) :

Tableau 3. Les consonnes du fulfulde

	labiales	apicales	palatales	postérieures
occlusives sourdes	*p*	*t*	*c*	*k*
occlusives sonores	*b*	*d*	*j*	*g*
glottalisées	*ɓ*	*ɗ*	*ƴ*	*ʔ*[3]
nasales	*m*	*n*	*ny*	*ŋ*
prénasalisées	*mb*	*nd*	*nj*	*ng*
fricatives sourdes	*f*	*s*		*h*
fricatives sonores	*v*	*z*		
continues	*w*	*l*	*y*	
battue		*r*		

2.4 Les voyelles du fulfulde

Le système vocalique peul s'étage sur trois degrés d'aperture et présente une opposition de quantité (Tourneux et al. 2021: 13). Il ne comporte pas de nasales phonologiques, mais le contexte *C_ŋ* provoque une nasalisation phonétique.

Tableau 4. Les voyelles brèves et longues du fulfulde

i	*u*	*ii*	*uu*
e	*o*	*ee*	*oo*
a		*aa*	

3 Cette consonne, très fréquente dans la langue, ne se transcrit pas quand elle est à l'initiale de mot. Dans les autres positions, elle s'écrit / ' /.

DOI 10.15460/auue.2025.98.1.332

2.5 Comparaison des deux systèmes

Du point de vue phonétique, principale différence entre les deux systèmes vocaliques : le fulfulde ne possède pas de voyelle centrale haute [ə].

Pour ce qui est du système consonantique, le fulfulde n'a pas de labio-vélaires *(kʷ, gʷ, ngʷ)* ni de latérales fricatives *(ɬ, ɮ)*. Le giziga, quant à lui, est dépourvu de nasale palatale *(ny)* et de glottalisée palatale *(y)*.

Le fulfulde est une langue sans tons, alors que le giziga a deux registres tonals (haut et bas). Nos sources giziga ne notent malheureusement pas les tons, sauf à de très rares exceptions.

2.6 Intégration phonologique des emprunts

Tout compte fait, l'intégration phonologique d'emprunts – peuls en giziga et giziga en peul – ne doit pas être très difficile, du fait de la proximité relative de leurs inventaires de phonèmes. Cela se vérifie effectivement. Cependant, les procédés d'intégration mis en œuvre de part et d'autre ne sont pas systématiques.

2.6.1 Dans le sens giziga > fulfulde

La consonne la plus problématique du giziga, à savoir la latérale fricative sonore, est remplacée en fulfulde par une simple latérale :

mbaɮa > *mbal* « bière de mil »

2.6.2 Dans le sens fulfulde > giziga

On observe parfois un phénomène (a) d'abrègement de la voyelle longue du fulfulde et (b) de dégémination consonantique.

2.6.2.1 Abrègement de la voyelle longue du fulfulde

Fulfulde		Giziga	
arngaawo	>	*arngawo*	« lit »
baaba	>	*baba*	« père »
baatal	>	*batal*	« aiguille, seringue »
boo	>	*bo*	« aussi, également »
buɗaare	>	*buɗare*	« vagabond »
jaɓɓaama	>	*jaɓama*	« bienvenue ! »

DOI 10.15460/auue.2025.98.1.332

maraawo	>	*marawo*	« gifle »
saakre	>	*sakre*	« pépinière (de mouskouari) »
saayiire	>	*sayre*	« bonnet en toile à matelas »

L'abrègement vocalique est assez aléatoire et l'on peut supposer que les Guiziga les plus familiers de la langue peule ont tendance à respecter la longueur vocalique du mot peul.

Fulfulde		Giziga	
caaca	>	*caca ~ caaca*	« jeu de hasard »
luumo	>	*lumo ~ luumo*	« marché, semaine »

2.6.2.2 Dégémination de la consonne du fulfulde

Le lexique giziga ne semble pas présenter de consonnes géminées, ce qui explique qu'il les réduit à l'unité dans les emprunts qu'il fait au fulfulde.

Fulfulde		Giziga	
callallu	>	*celelew*	« chaîne »
cette	>	*cete*	« brochette de viande »
jaɓɓaama	>	*jaɓama*	« bienvenue ! »
luggere	>	*lugere*	« bonne terre, terre fertile »

2.7 Intégration morphologique des emprunts giziga en fulfulde

La marque la plus voyante d'intégration d'un lexème giziga en fulfulde est l'attribution d'un suffixe de classe, ou classificateur. En voici quelques exemples :

	Giziga		Fulfulde	
-ki	*ngalalay*	>	*ngalalayhi*	« *Dalbergia melanoxylon* »
	maliya	>	*maliyaahi*	« *Ficus thonningii* »
	mindek	>	*mendekhi*	« *Ficus dicranostyla* »
	turus	>	*tursuuhi*	« *Haematostaphis b.* »
-ndi	*muzuk*	>	*muzukri*	« sorgho rouge sp. »
-ko	*mangaraw*	>	*mangaramho*	« *Corchorus tridens* »
	mbere-mbere	>	*mbere-mbereeho*	« *Commelina benghalensis* »

DOI 10.15460/auue.2025.98.1.332

	Giziga		Fulfulde	
-ndu	*kucum*	>	*kucumru*	« daman »
	petengew	>	*petengewru*	« petit batracien sp. »
-nde	*muray*	>	*murayre*	« *Brycinus nurse* »
	menjeɗe	>	*menjeɗeere*	« *Brycinus macrolepidotus* »
	gulek	>	*gulekre*	« hache à soie »
	jiyku	>	*ziikuure*	« bonnet en toile à matelas »
-ɗo	*bodogor*	>	*bodogorjo*	« célibataire »
	gudi	>	*gudiijo*	« forgeron »
	gula	>	*agulaajo*	« gaucher »

Certains lexèmes giziga sont cependant entrés en fulfulde sans recevoir de suffixe de classe :

Giziga		Fulfulde	
ardedel	>	*ardadel*	« *Merremia emarginata* »
gawla[4]	>	*gawla*	« porteur »
mbaɣa	>	*mbal*	« bière de mil »
murla	>	*murla*	« colostrum »
zindirɗa	>	*zindirɗa*	« ver de terre »

3 Inventaire des mots giziga qui sont passés en fulfulde

Nos ouvrages de référence pour le lexique fulfulde sont Tourneux & Yaya Daïrou (2017) et Parietti & Tourneux (2018).

Les Peuls arrivant dans un milieu naturel et social qu'ils ne maîtrisaient pas ont emprunté au giziga des noms de végétaux et d'animaux. On peut aussi présumer qu'ils connaissaient certaines de ces réalités mais que, les Guiziga leur accordant une importance particulière, c'est leur dénomination qui l'a emporté. Nous donnons ci-dessous les mots giziga qui sont entrés en fulfulde et nous les classons par domaines sémantiques.

4 Il y a un changement de sens entre le giziga *gawla* « jeune homme » et le fulfulde *gawla* « porteur ».

DOI 10.15460/auue.2025.98.1.332

3.1 Végétaux

Nous avons identifié les végétaux en giziga grâce aux travaux de Seghieri (sans date, et 1990) et Seghieri & Floret (1993). Plusieurs de ces végétaux sont utilisés en alimentation humaine (brèdes : *mangaraw, memel, mindek, mbere-mbere* ; fruitiers sauvages : *maliya, turus* ; variété locale de sorgho pluvial : *muzuk*). Le *Dalbergia* ou ébène des Égyptiens, est un excellent bois de feu et sert aussi à fabriquer des stylets et des massues (Tourneux & Yaya Daïrou 2017: 511). Le *Ficus thonningii* est aussi un arbre fourrager et il possède de nombreux usages médicinaux (Tourneux & Yaya Daïrou 2017: 374–375). Deux de ces plantes (*ardedel, mbere-mbere*) sont des adventices des cultures.

ardedel : « *Merremia emarginata* » (Convolvulaceae) ; ffde *ardadel* « *Merremia emarginata* ».

maliya : « *Ficus thonningii* (Moraceae) » ; > ffde *maliyaahi* « *Ficus thonningii* ».

mangaraw : « *Corchorus tridens* (Tiliaceae) » ; > ffde *laalo-mangaramho* « *Corchorus tridens* » cf. mofu-Gudur *mangəráw* « herbe à sauce, que l'on mélange avec les haricots concassés » (Barreteau 1988: 181).

mbere-mbere : « herbe rampante » ; **mbere-mbere maja** : « *Commelina benghalensis* » ; > ffde *mbere-mbereeho* « *Aneleima lanceolatum* et *Commelina benghalensis* (Commelinaceae) » ; cf. mofu-Gudur *mémbéré-mbere* « Commélinacée sp. ».

memel ~ memeɗ : « herbe à sauce très gluante » > ffde *memelho, memeyelho* « *Corchorus fascicularis* (Tiliaceae) » ; cf. mafa *málámá* « herbe à sauce sp. » (Barreteau & Le Bléis 1991).

mindek : « Ficus sp. » ; > ffde *mendekhi* « *Ficus dicranostyla* (Moraceae) » ; cf. mofu *məndek*, mafa *mindek* (Barreteau & Le Bléis 1991).

muzuk : « variété de sorgho rouge pluvial » ; cf. ffde *mujukri* « variété de sorgho », de l'ethnonyme *Muzuk / Mujuk / Munjuk*.

ngalalay : « *Dalbergia melanoxylon* » ; > ffde *ngalalayhi* « *Dalbergia melanoxylon* (Fabaceae) ».

turus : « *Haematostaphis barteri* » > ffde *tursuuhi* « *Haematostaphis barteri* (Anacardiaceae) ».

3.2 Animaux

Les poissons cités dans la liste ci-dessous *(malalay, maray, menjeɗe, mepelele, ndola)* avaient et ont encore une importance économique reconnue dans la région, vu leur abondance dans les cours d'eau et les innombrables mares (Blache et al. 1964). Ceux que l'on trouve actuellement sur les marchés, frais ou secs, proviennent plutôt de

DOI 10.15460/auue.2025.98.1.332

lacs (Maga) ou des fleuves Chari et Logone, ainsi que des yaérés (plaines inondables).

En revanche, les batraciens comestibles *(merlek, petengew)* restent une spécialité strictement locale (Seignobos 2014). Ces animaux constituaient manifestement une nouveauté pour les Peuls venant de l'Ouest, qui n'avaient pas jusqu'alors l'idée de les consommer. Notons qu'il existe encore localement d'autres batraciens comestibles, mais dont les Peuls n'ont pas emprunté le nom au giziga. C'est le cas principalement de la grenouille-taureau *(gidigidiiru) Pyxicephalus adspersus*, dont le nom peul proviendrait du gidar (kaɗa) (Seignobos 2014: 308).

Le daman *(kucum)* fournit une chair appréciée des populations habitant au pied des collines rocheuses.

Pour conforter le sens de l'emprunt, nous indiquons ci-dessous des cognats tchadiques, prouvant que c'est bien le fulfulde qui est la langue emprunteuse.

kucum : « daman des rochers » ; > ffde *kucumru* daman, *Procavia capensis* (Procaviidae) ; ce mot appartient bien au tchadique : cf. mofu-Gudur *kucam* « daman des rochers ; munjuk *kusum* « rat, souris ».

malalay : « poisson sp. très glissant, avec des nageoires piquantes » ; > ffde *malalaywu* « *Bagrus bayad* et *B. docmak* » (Bagridae) ; cf. mofu-Gudur *malalay* « poisson sp. ».

maray ~ muray : « poisson à queue à moitié rouge qui ressemble à celle de la "sardine" » ; > ffde *murayre* « *Brycinus nurse* (Alestidae) [= *Alestes nurse* (Rüp. 1832), (Characidae)], "sardine" ». Le mofu-Gudur *mbəráŋ* « sardine », montre qu'on a bien affaire à un mot tchadique.

mbirlek ~ mburleke : « grenouille *Ptychadena trinodis » ; ffde merlekru*, cf. mofu-Gudur *márlek* « grenouille sp. », plus proche du ffde que le giziga.

menjeɗe ~ manjaɗay : « petit poisson gras » ; > ffde *menjeɗeere* « *Brycinus macrolepidotus* [= *Alestes macrolepidotus* (C.V. 1869), (Characidae)] (Characidae) » ; cf. mofu-Gudur *máanjaɗáy* « alevin ».

mepelepele : « poisson à chair rouge, gras et plein d'arêtes » ; > ffde *pelpeleewu* « *Alestes dentex* et *A. baremoze* (Alestidae) ».

ndola : « poisson noir allongé à petites écailles [d'après Michielan ; en fait, ce poisson est sans écailles] » ; > ffde *ndoolaawu* « *Mormyrus rume* (Mormyridae) » ; Barreteau (1988: 205) donne, à tort probablement, le mot *ndáwla* « poisson sp. » comme « emprunté au ffde ». En effet, le mot correspond bien à une racine tchadique ; voir kotoko de Goulfey *dòlá* (**dàwlá*), kotoko de Makari *tólá* (**táwlá*), kotoko de Kousseri *ǹdòlá* (**ǹdàwlá*) (Tourneux 2013: 192).

DOI 10.15460/auue.2025.98.1.332

petengew : « petit batracien qui gonfle le ventre si on le touche (Seignobos 2014: 308), *Hemisus marmoratus sudanensis* » ; > ffde *petengewru* « petit batracien qui passe pour faire crever les animaux qui l'avalent » ; cf. mofu-Gudur *pétéŋgew* « rainette ».

zindirɗa : « ver de terre » : > ffde *zindirɗa* « ver de terre » ; le mot est bien d'origine tchadique ; voir mofu-Gudur : *mázərɗáɗa* « ver de terre », mot dérivé du verbe *zərɗaɗaɗa-ɗá* « s'étirer en rampant », lui-même dérivé de *-zárɗ-* « s'étirer » (Barreteau 1988: 264).

3.3 Culture matérielle

Les Peuls ont aussi emprunté des vocables ayant trait à des activités techniques : agriculture *(gulek)*, cuisine et alimentation *(mbaɮa, mogoyok)* habillement *(jiyku)*.

gulek : « pioche » ; > ffde *gulekre* « hache à soie » ; *gulekwo* « houe à soie » ; *gulekyel* « petit fer plat à douille de l'outil appelé *gasirgal* » (Tourneux 1984) ; cf. mofu-Gudur *gùlèk* « houe longue et pointue ».

jiyku : « bonnet » ; > ffde *ziikuure* « bonnet en toile » ; le mot est bien d'origine tchadique ; cf. mofu-Gudur *jiygaw, jiykwaw* « bonnet » ; wandala *dzakwa* « bonnet ».

mbaɮa : « bière de mil » ; > ffde *mbal* « bière de mil » ; cf. mofu-Gudur *mbaɮa* « bière de mil chaude » et nombreuses autres langues tchadiques.

mogoyok : « cendres de tiges de mil » ; > ffde *mogoyokri* « saumure végétale résultant de la lixiviation de cendres de tiges de mil pénicillaire ».

3.4 Structure socio-politique

bodogor : « célibataire » > ffde *badigorjo* ~ *bodogorjo* « célibataire, homme dont la femme est absente, veuf ».

bugawla : [litt : « chef jeune homme »], « vainqueur, le plus fort ; héros » ; > *ffde bigawla* « esclave de confiance du chef ».

gáwla : « jeune homme » ; > ffde *gawla* « porteur ».

gudi : « forgeron » ; > ffde *gudiijo* « forgeron » (la comparaison avec ffde du Mali *baylo* démontre que le ffde du Diamaré *gudiijo* a été emprunté) ; voir Tourneux 1991.

3.5 Autres

gula : « gauche ; gaucher » ; > ffde *agulaajo* « gaucher, gauchère » (cette racine est bien tchadique ; cf. mofu-Gudur *gula* « gauche »).

murla ~ mərla ~ mirla : « colostrum » ; > ffde *murla* « mauvais lait maternel, que les femmes identifient au colostrum » (Tourneux et al. 2007: 311).

DOI 10.15460/auue.2025.98.1.332

4 Inventaire des mots giziga venus du fulfulde

Les mots giziga suivants viennent du fulfulde (nos ouvrages de référence pour le lexique fulfulde sont Noye [1989] et Parietti & Tourneux [2018]). Nous les classons avec quelque arbitraire en sept domaines sémantiques.

4.1 Végétaux

ɓokki : « baobab » ; du ffde *ɓokki* « baobab ». Cet emprunt fait double emploi en giziga, qui a le mot *mulguy* ou *mulugwiy* pour désigner cet arbre.

kumambeɗe : « herbe sp. » ; du ffde *kummbo-mbetta* « *Dactyloctenium aegyptium* (Poaceae) ». Le sens littéral des composants du mot peul est : « Kummbo qui reste à plat ». Cela fait référence au fait que la plante en question est une Graminée à base couchée à rampante. « Kummbo » est un anthroponyme peul féminin qui réfère à l'ordre de naissance (deuxième fille). En giziga, il n'est pas possible de décomposer *kumambeɗe* en éléments signifiants.

4.2 Animaux

awra : « âne sp. », du ffde *awraare* « âne à pelage gris clair » ; cf. ffde du Mali *araawa* « âne ».

caɗ(u)ngu : « animal mort sans avoir été égorgé » ; du ffde *caaɗngu* « animal crevé ».

kocori ~ kocoro ~ kocorop ~ hosori ~ husuru : « 1. escargot ; 2. coquille d'escargot ; via ffde *hoosooru / koosooji* « coquille d'escargot » ; cf. mofu-Gudur *kwécérew*, mafa *kwecéré* « coquille d'escargot ; toupie » (donnés par Barreteau [1988] et Barreteau & Le Bléis [1991] comme emprunts au ffde) ; *hosooru* est attesté au Foûta Tôro et au Niger occidental (De Wolf 1995, vol. 3, « S »: 219) ; cf. perge tegu (dialecte dogon) *kŏwsò* « coquille d'escargot » (Segerer & Flavier 2011–2019).

maaba : « coucal du Sénégal », syn. *morogoyogoyo* ; du ffde (*colla-maabani* « coucal du Sénégal », cf. *maaboo* « chanter les louanges de » ; l'oiseau en question a un chant très bruyant.

molde : « ânon » ; du ffde *molde* « ânon » (du radical *mol-* « petit [d'un animal] ») ; ffde du Mali *mola* « ânon, chamelon… ».

ngaari : « taurillon » ; du ffde *ngaari* « taureau ».

4.3 Culture matérielle

Nous remarquons notamment deux éléments d'ameublement *(arngawo, karawal)*, de vêture *(horende, mbolori, metalel, sayre)* un élément architectural *(ɓamtal)* dont les noms ont été empruntés à la

DOI 10.15460/auue.2025.98.1.332

langue peule en même temps que l'élément matériel lui-même. Traditionnellement, l'habitation circulaire giziga était de faible diamètre ; son toit autoportant ne nécessitait donc pas l'emploi d'un poteau central ; sur le modèle de la case peule, l'architecture giziga s'est modifiée, en construisant des maisons d'un diamètre beaucoup plus important, nécessitant, cette fois, l'usage du poteau central (C. Seignobos, c.p., 04.05.2023).

arngawo ~ (h)irngawo : « lit » ; du ffde *arngaawo* « lit ».

ɓamtal : « poteau central » ; du ffde *ɓaŋtal* « poteau central destiné à étayer la toiture ».

batal : « aiguille ; seringue ; injection », du ffde *baatal* « aiguille, seringue, injection » ; cf. ffde du Mali *baatal* « aiguille à coudre ; aiguille à injections ».

bohal ~ bohol ~ bowal ~ buwal ~ bowagol : « grand-route » ; du ffde *buuwol* « grand-route » ; cf. ffde du Mali *buuwaangol* « grande route ».

celelew : « chaîne en fer » ; du ffde *callallu* ; cf. ffde du Mali *callalol* « fil métallique, chaîne ».

cete : « brochette de viande » ; du ffde *cette* (plur.) « brochettes de viande » ; ffde du Mali *cettal/cette* « brochette (nue ou garnie) ».

darogal ~ dorogal : « miroir, glace » ; du ffde *daaroogal ~ daarorgal* « glace, vitre ».

ɗerewol ~ ɗerewel : « papier ; livre, cahier » ; du ffde *ɗereewol* « feuille de papier ».

horende : « chéchia de couleur rouge » ; du ffde *woronnde* « chéchia » (il y a peut être, en giziga, contamination entre deux mots peuls : *woronnde* « chéchia » et *hoore* « tête ».

karawal : « siège, chaise » ; du ffde *koromwal ~ koroowal* « banc, chaise, fauteuil, siège ».

mbolori ~ mbulori : « chapeau de paille » ; du ffde *mbuuloore* « chapeau de paille à larges bords ».

metalel : « turban » ; du ffde *meetaleewol* « turban », ffde du Mali *meetalol* « bande de tissu ; turban ».

mohol : « mur d'enceinte » ; du ffde *mahol* « mur » (dérivé du verbe *mah-a*, 1. « façonner (une poterie) » ; 2. « construire »).

sayre : « bonnet en toile à matelas » ; du ffde *saayiire* « bonnet en tissu gris foncé (avec rayures) » ; cf. hausa *ɗan shàyíi* « garçon incirconcis »

ta(a)ri : « fronde en caoutchouc » ; < ffde *taari* « 1. cire d'abeille ; 2. gomme végétale caoutchouteuse » ; cf. ffde *taari* « 1. cire ; 2. latex, caoutchouc ».

DOI 10.15460/auue.2025.98.1.332

4.4 Agriculture et élevage

bortol : « chemin par où passe le bétail » ; du ffde *burtol* « piste à bétail » ; ffde du Mali *buurtol* ; cf emprunt du hausa au ffde *búrtalíi*.

jobol : « ligne, rang, rangée » ; du ffde *jabbol* « ligne de culture ».

karal : « sol argileux pour sorgho de contre-saison » ; du kanuri [hár-] « être sec » d'après Mohammadou Eldridge (1997). Cependant, la consultation du *Dictionnaire peul-français (Mali)* de C. Seydou remet en cause cette origine kanuri. On y trouve en effet *karal* « étendue de terre sèche, dure et nue (généralement latéritique) ».

lugere : « bonne terre, terre fertile » ; du ffde *luggere* « bas-fond argileux, dépression de terrain ».

saɓɓa : « paille, herbe qu'on coupe pour aller la mettre dans le champ de karal [terrain à sorgho repiqué] avant d'y mettre le feu » ; du ffde *saɓɓa* « étaler » cf. l'expression *saɓɓa karal* « étaler l'herbe qu'on a coupée sur le champ de saison sèche (pour ensuite y mettre le feu) » (Tourneux & Yaya Daïrou 2017: 586).

sakre : « pépinière (en particulier de sorgho de contre-saison) » ; du ffde *saakre* « pépinière de sorgho à repiquer » (du radical verbal *saak-a* « semer à la volée ») (Tourneux & Yaya Daïrou 2017: 582).

4.5 Pratiques et institutions sociales

baba : « père » ; du ffde *baaba* ou du kanuri *babá* « père ».

bambaɗo : « griot » ; du ffde *bammbaaɗo* « griot ».

ɓesde : « ajout, cadeau qu'on ajoute sur ce qu'on a acheté » ; du ffde *ɓesda* « ajouter », *ɓesdee* « ajoutez ! ».

buɗare : « vagabond » ; du ffde *buuɗaare* « voyou, vagabond ».

caca ~ caaca : « jeu de hasard » ; du ffde *caaca* « jeu de hasard », ffde du Mali *caaca* « jeu de cauris, jeu de hasard », cf. hausa *cáacà* « jeux d'argent ».

dimu : « liberté » ; du ffde *ndimu* « liberté ».

giɗaɗo : « épouse courageuse au travail » ; du ffde *giɗaaɗo* « (personne) aimée ».

jaɓama : « bienvenu ; bienvenue ! » ; du ffde *jaɓɓaama* ! « bienvenue ! ».

luma ~ lumo [~ luumo, Shay 2021: 372] : « marché ; semaine » ; du ffde *luumo* « marché ».

marawo : « gifle » ; du ffde *maraawo* ; cf. ffde du Mali *marhaaniiwo*, *mbaraawo* ; le mofu-Gudur a *máaráwa* « gifle », que D. Barreteau (1988: 183) signale comme un emprunt au ffde.

meɓɓere : « foule » ; du ffde *mooɓre* « foule, rassemblement ».

nara : « entente » ; du ffde *narra* « être d'accord, s'entendre ».

DOI 10.15460/auue.2025.98.1.332

ngorgi : « même classe d'âge, classe d'initiation » ; < ffde *ngorgi* « camarade de même âge, circoncis la même année » ; cf. ffde *ngorgiɗa* « être du même âge » (Taylor 1932: 67).

4.6 Lexèmes relatifs à la perception

deno-deno : « bleu » ; du ffde *deenoo-* (adj.) « bleu clair ».
dus : « odeur nauséabonde » ; de l'adverbe idéophonique ffde *dus* « (sentir) très mauvais » ; cf. ffde du Mali *duss* « (sentir) très mauvais » [Seydou sous presse, sous *borde* « crottes »].
jawal : « rapidité, précipitation » ; du ffde *jaawal* « rapidité, vitesse » (dérivé de la racine verbale *yaaw-a* « être rapide ; aller vite »).

4.7 Adverbes et éléments grammaticaux

bo : aussi, également » ; du ffde *boo* « aussi, également ».
caka cak : « juste par le milieu, juste au milieu » ; du ffde *caka cak* « en plein milieu ».
kurum : « très noir » ; du ffde du Diamaré et ffde du Mali *kurum* très (noir) ; cf. baguirmien (Nilo-Saharien) *cɬə́m* « très (noir) ».
malla : « ou bien » ; du ffde *malla* « ou bien ».
njáɓ : « idéophone : qui se conviennent physiquement ou moralement » ; du ffde *jaɓa* « accepter, admettre, agréer » (au pluriel *njaɓa*).
siriw : « silencieux » ; du ffde *siriw* « (idéophone) sans rien dire, en silence ».
siké : « alors, vraiment » ; du ffde *sika* « particule énonciative interrogeant sur la véracité de l'énoncé ».
tal tal : « tout blanc (d'une certaine étendue) ; < ffde *tal* « intensificateur, très (blanc, propre) » ; cf. ffde du Mali *tal* « très (blanc, propre) ».

5 Le fulfulde comme plate-forme d'échanges linguistiques

Le fulfulde a servi au giziga de plate-forme sur laquelle il est allé emprunter à de nombreuses autres langues – allemand, anglais, arabe, bambara, hausa, kanuri, lingua franca, sango, seereer, soŋay, soninke, tchadique – parfois via plusieurs intermédiaires. Il n'est donc pas toujours facile de déterminer l'origine ultime d'un mot. Le hausa lui-même, par exemple, a emprunté à de multiples sources (Skinner 1996) auxquelles nous tenterons de remonter. Dans tous les cas, nous tiendrons compte de la vraisemblance phonétique, sémantique et historique pour nous assurer de la direction de l'emprunt.

DOI 10.15460/auue.2025.98.1.332

5.1 Emprunt du giziga à l'allemand via le hausa et le fulfulde

dala : « argent ; unité monétaire valant 5 francs » ; de l'allemand *Thaler*, via le hausa *dálàa* « 1. thaler de Marie-Thérèse ; 2. ancienne pièce de deux shillings ; 3. ancienne pièce de 5 francs » ; via ffde *dala* « pièce de 5 francs ».

5.2 Emprunts du giziga à l'anglais via le hausa et le fulfulde

bariki : « bureau, administration », de l'anglais *barracks*, via le ffde *baariki* « bureau, bureau administratif » ou le kanuri *barîkí*, qui sont passés eux-mêmes par le hausa *báarìkì* « campement, caserne ».

is : « levure » ; de l'anglais *yeast* « levure », via ffde *iis* « levure chimique » et le hausa *yîs*.

kalanjir ~ kalanjur : « lampe en terre fonctionnant avec de la résine » ; de l'anglais *kerosene* [ˈkerəsiːn] dans *kerosene lamp* « lampe à pétrole », via hausa *kànànzîr* « pétrole (lampant) » et ffde *kalanjir* « pétrole ».

koɓo : « petite pièce de monnaie » ; du hausa *kwábòo* « penny, ancienne monnaie du Nigeria » ; via ffde *koɓo* « petite pièce de monnaie » ; cf. kanuri *kóbo* « penny, ancienne monnaie du Nigeria » ; d'après Skinner (1996: 158) le mot viendrait de l'anglais *copper*, qui, en argot signifie « penny ».

lammba : « numéro ; marque » ; de l'anglais *number* ; via ffde *lammba* « numéro ; marque », via hausa *lámbàa* « signe, marque distinctive ; numéro minéralogique » ; cf. kanuri *lámba* « nombre, numéro d'immatriculation ».

sisi : « ancienne monnaie » ; de l'anglais *six* (Skinner 1996: 233), via hausa *sîisìi* « pièce de sixpence (ancienne monnaie du Nigeria), via ffde *siisi* « pièce de 5 francs ».

soje : « soldat, gendarme, policier » ; de l'anglais *soldier* [ˈsəʊldʒəʳ] ; via ffde *sooje* « militaire, policier, homme en tenue », via hausa *sóojà* « soldat, militaire » et kanuri *sója* « soldat ».

suloy : « ancienne monnaie » ; de l'anglais *shilling*, via ffde *suleyre* « ancienne pièce de monnaie (10 francs) », via le yoruba *ṣílè* [ʃile], > hausa *súlàì* « shilling », ancienne monnaie du Nigeria.

tastilam : « lampe torche » ; de l'anglais *torch-lamp*, via ffde *tostilam* « lampe torche » et hausa *tóocìlàn* « lampe torche ».

5.3 Emprunts du giziga à l'arabe, via le fulfulde

Le fulfulde a largement emprunté à la langue arabe (Lacroix 1967), cela en fait donc un intermédiaire privilégié entre l'arabe et le giziga. Nos ouvrages de référence pour le lexique arabe sont Jullien de Pommerol (1999) et Baldi (2008).

DOI 10.15460/auue.2025.98.1.332

abada : « pagne de piètre qualité » ; de l'arabe *abadan* « pour toujours », via ffde *abadaare* « pagne dont les teintes ne passent pas ; notons l'important changement de sens que le mot a subi en passant du fulfulde au giziga et la disparition du suffixe de classe.

abada : « toujours » ; de l'arabe *abadan* « pour toujours », via ffde *abada* « toujours ».

akre : « rétribution en nature ; location » ; de l'arabe *kirā'* « louer », via ffde *akiri* « indemnité en nature ou en espèces, redevance ; location ».

amma : « mais » ; de l'arabe *ammā*, via ffde *ammaa* ou kanuri *ammá*.

arge ~ erge : « alcool » ; de l'arabe *araq*, via arabe tchadien *argi* et ffde *arge* « alcool artisanal ».

asar : « malheur, accident, situation triste » ; de l'arabe *kasāra* « perte, dommage », via ffde *asar* « dépense, perte, ruine » ou kanuri *asár* « perte ».

borgo : « couverture » ; de l'arabe *burqu'* « voile », via ffde *borgo* « couverture » ; cf. kanuri *borkó* « couverture » ; cf. ffde du Mali *borgoore* « couverture blanche tissée ».

dabaray : « moyen, ruse, stratégie, tromperie » ; de l'arabe *dabbara* « projeter, dresser un plan », via ffde *dabare* « manière, moyen, stratagème » ; cf. hausa *dàbáaràa* « plan, idée, ingéniosité ».

dangay : « prison » ; de l'arabe tchadien *dangay* « maison en terre, chambre, prison » ; via ffde *saare danngaay* « prison » ; cf. hausa *dángáa* « clôture en tiges de mil » et kanuri *dángá* « clôture en tiges de mil ».

dawra : « gandoura » ; de l'arabe *qandūr* « coquet, pimpant », via kanuri *dáurá* et ffde *dawrawol* ; cf. ffde du Mali *ganduuru*.

dole : « obligatoire, obligé » ; de l'arabe *daula* « pouvoir », via hausa *dóolè* « nécessairement » > ffde *doole* « obligatoirement, par force ; cf. wolof *doole* « force, puissance ».

dulniya ~ duniya : « 1. monde ; 2. la vie, les gens » ; de l'arabe *duniyā* « monde » via ffde *duniya* « bas monde, terre, vie sur terre ».

fakat : « exactement, sûrement, certainement » ; de l'arabe *faqaṭ* « seulement », via ffde *fakat* « assurément ».

godo : « couverture blanche en coton tissée à la main » ; de l'arabe *quṭn* « coton », via ffde *godoore* « bande de coton tissée à la main ».

hàràm : « interdit, refus catégorique » ; de l'arabe *ḥaram* « interdit », via ffde *haram* « interdit par la religion ».

jam : « exclamation de contentement, d'agrément » ; cf. **jamjam.**

jamjam : « en paix, en bonne santé » ; de l'arabe *jamma* « se reposer », via ffde *jam* « paix, bien-être, tranquillité ».

jiyba : « poche », de l'arabe *jībat* « poche », via ffde *jiiba* « poche de vêtement ».

labara : « nouvelle » ; de l'arabe *kabar* (précédé de l'article *al*) ; via hausa *làbarìi* « nouvelle(s), information » ou kanuri *lawâr* (**labâr*) ; cf. kanuri

DOI 10.15460/auue.2025.98.1.332

hawâr (**habar*) et ffde *habar, habaru* (qui n'ont pas intégré l'article arabe).

laya : « amulette, grigri » ; de l'arabe *āya* (précédé de l'article *al*) « symbole, marque ; verset coranique » via kanuri *láya* « amulette, charme » ; cf. ffde *layaaru* « amulette, grigri ».

lemu : « citron » ; de l'arabe *laimūn* « citron lime », via ffde *leemu* « citron ».

mekef : « ciseaux » ; de l'arabe *miqaṣṣ* « ciseaux » ; via ffde *mekefje*.

nafa : « utilité, intérêt » ; de l'arabe *naf'* « avantage, profit » ; via ffde *nafa* « être utile ».

pitirla : « lampe à pétrole » ; de l'arabe *fatīla* « 1. corde roulée ; 2. mèche, de lampe ou de bougie » (Baldi 2008: 375) via ffde *pittirla* ~ *fittirla* « lampe à pétrole » ; cf. hausa *fìtílàa* « lampe », yoruba *fitilà* « lampe à huile » ; on ne peut exclure une contamination avec l'anglais *petrol* dont *pitirla* comporte, dans l'ordre, toutes les consonnes ; dans la langue contemporaine, *petrol* signifie « essence », certes, mais ce n'était pas le cas autrefois, quand on ne distillait pas le pétrole.

riba : « gain, bénéfice » ; de l'arabe *ribḥ*, plur. *arbāḥ* « bénéfice, intérêt », via ffde *riba* « bénéfice, gain » ; cf. kanuri *ríwa* (**ríba*) « profit, intérêt, gain ».

sa'e : « 1. café ou thé ; 2. lie de bière fermentée » ; de l'arabe *šāy* « thé » (emprunt à une langue du sud-est asiatique), via ffde *saa'i* « thé » ; cf. hausa *sháayì* « thé ».

sariya : « jugement » ; de l'arabe *šari'a* « loi islamique », via ffde *sariya* « loi islamique ; sentence ».

sedege : « aumône faite en sacrifice » ; de l'arabe *ṣadaqa* « aumône », via ffde *sadaka* « aumône, offrande, sacrifice ».

sedewo : « témoin » ; de l'arabe *šāhid* « témoin », via ffde *ceedoowo* « témoin » (dérivé du radical *seed-oo* « témoigner »).

sirla : « pantalon » ; de l'arabe *sirwāl* « pantalon », via ffde *sirla* « pantalon ».

sukwar (dans *mandaŋ sukwar*) : « sucre » : de l'arabe *sukkar* « sucre, via ffde *suk(k)ar* « sucre.

tambal : « grand tambour que l'on trouve chez les lamibé » ; de l'arabe *ṭabl* « tambour » (Baldi 2008: 312 ; Erlmann 1983: 19) ; via kanuri *tə́mbal* « grand tambour en bois suspendu à l'horizontale entre deux supports » et ffde *tummbal* « grand tambour cylindrique à deux membranes ».

urdi : « parfum » ; de l'arabe *'iṭr* / plur. *'uṭur* (avec métathèse consonantique > *'-r-ṭ*), via ffde *urdi* « parfum ».

yáwwa(a) : « exclamation de contentement » ; de l'arabe tchadien *yawwâ* « oui !, d'accord ! » (< arabe *aywā* « oui »), via hausa *yâwwáa* ! « réponse à une salutation ; 2. c'est juste ! ; 3. bravo ! » ; cf. ffde *yowwaa* ! « interjection marquant le contentement, l'acquiescement ».

DOI 10.15460/auue.2025.98.1.332

zaaman : « autrefois » (étymologie populaire : « au temps des Allemands ») ; de l'arabe *zamān* « autrefois ; temps jadis » ; via ffde *zamaanu* « époque, siècle », via hausa *záamànîi* ~ *zàamáanìi* « période, époque », plus contamination avec ffde *Jaaman(jo)* « Allemand ».

zakka : « dîme, redevance en nature versée aux chefs » ; de l'arabe *zakāh* « aumône obligatoire » ; via hausa *zàkkáa* « dîme » et ffde *zakka* « dîme » ; cf. kanuri *zaká* « aumône obligatoire ».

5.4 Emprunts du giziga au bambara via le fulfulde

Notre ouvrage de référence, pour le bambara, est le dictionnaire de Gérard Dumestre (2011).

goŋgoŋ : « boîte ou fût métallique » ; du bambara *góngon* « touque » via ffde *goŋgoŋ* ; cf. ffde du Mali *gonngooru* « bidon, fût métallique », cf. hausa *gwángwání* « boîte en fer-blanc ».

senge : « moustiquaire » ; du bambara *sánge* ~ *sánke* « palissade, rideau, clôture, moustiquaire » (Gouffé 1971: 59–61), via hausa du Niger *sangee* et ffde *sange* « moustiquaire ».

5.5 Emprunts du giziga au hausa via le fulfulde

Nos ouvrages de référence pour le lexique hausa sont Bargery (1993), Newman (2007), Newman & Ma (1982, 2020).

adiku : « foulard, mouchoir de tête » ; du hausa *àdíikòo* « fichu, foulard de femme », via ffde *adikko*, *adiiko* « foulard, mouchoir de tête ».

alawaya : « tunique d'homme à manches courtes, descendant jusqu'aux genoux » ; du hausa *áláwáyyòo* « vêtement en calicot blanc », via ffde *alawoya* « calicot, tissu blanc ».

asusu : « caisse à argent », du hausa *ásúusù* « tirelire », via ffde de Garoua *asuusu* « tirelire ».

barasa : « alcool, eau de vie » ; du hausa *bàaràasáa* « toute boisson fortement alcoolisée », via ffde *baaraasa* « alcool, boisson alcoolique » ; d'après Skinner (1996: 16), le mot vient du nom de l'île de Brass (Nigeria), par où l'on importait du gin au 19e siècle.

circir : « très droit » ; du hausa *cír* « tout droit », via ffde *cir* « tout droit, directement ».

dakare : « qui a une mauvaise conduite ; individu qui a une mauvaise conduite » ; du hausa *dàakáarè* « 1. fantassin ; 2. personne qui ose dire ce que d'autres considèrent comme inconvenant (Bargery 1993) » ; via ffde *daakaare* « personne de mauvaises mœurs » ; d'après Skinner (1996: 41), viendrait du tuareg *dakare* « troupe de fantassins », mais l'auteur ne cite pas sa source.

DOI 10.15460/auue.2025.98.1.332

dankali : « patate douce » ; du hausa *dànkálìi* « patate douce », via ffde *dankali* « patate douce ».

danko : « fronde en caoutchouc » ; du hausa *dánƙòo* « gomme, caoutchouc ; fronde » via kanuri *dánko* « caoutchouc » et ffde *danko* « caoutchouc ».

dayday ~ deydey : « moyennement ; de taille moyenne » ; du hausa *dáidái* « correctement, exactement », via ffde *deydey* « exactement ; quantité exacte ».

diga : « depuis » ; du hausa *dàgà* via ffde *diga*.

gajeré : « culotte » ; du hausa *gàjéeɽée* « court ; short », via ffde *gajeere* « culotte courte, short ».

garama : « impôt, taxe » ; du hausa de l'Adamawa (Taylor 1932) « taxes douanières » ; ffde *garaama* « impôt personnel, taxe ».

garaya : « "guitare", luth » ; du hausa *gàɽáayáa* « luth à deux cordes », via ffde *garayaaru* « luth des non-Peuls ».

gidaŋmucu : « cimetière » ; du hausa *gídán mútù*, litt. « concession de mourir » (l'expression n'est pas attestée dans les dictionnaires hausa standard [Sergio Baldi, courriel du 09.01.2021]) ; via ffde *gidammutu* « cimetière ».

iriŋ : « qualité ; semblable » ; du hausa *íɽìi* + *-n* (marque de détermination) « sorte, espèce », *íɽìn* « comme » ; via ffde *irin* « sorte, espèce ».

jawjaw : « tambour d'aisselle » ; du hausa *jáujé* « tambour semblable à *kalaŋgu* "tambour sablier", mais plus grand » (Bargery 1993: 494 et Newman 2007: 98), via ffde *jawjawru* « tambour d'aisselle » (Taylor 1932: 98).

jawleeru : « vestibule, porte d'entrée » ; du hausa *záuɽèe* « pièce par laquelle on entre dans une concession », via ffde *jawleeru* « case-vestibule ».

kay ! kayya ! : « interjection marquant l'étonnement, la surprise, le mécontentement » ; du hausa *kâi* ! « exclamation employée pour exprimer une légère désapprobation, un doute, une surprise » ; via ffde *kay* ! « interjection marquant l'étonnement, la surprise, le mécontentement ».

kose : « sorte de beignet » ; du hausa *ƙóosái* « beignet de farine de haricots » ; via ffde *koosay* « beignet de farine de niébés ».

kotóróko ~ kotórko : « radier, pont » ; du hausa *kàdáɽkò* « pont » (Bargery 1993: 520, revu par Newman (2007: 102) *kàdáɽkòo* « petit pont », et Baldi (2015: 157 : *kàdáɽkòo* « petit pont, temporaire), via ffde *katarko* « radier, pont » (Taylor 1932 attribue une origine hausa à ce mot sans fournir d'étymon).

kunu : « boisson à base de farine de mil, légèrement alcoolisée » ; du hausa *kùnúu* « bouillie de farine épicée » ; via ffde *kunu* « boisson à base de farine de mil, légèrement fermentée » ; cf. kanuri *kúndu* « sorte de bouillie préparée à base de jujubes ».

DOI 10.15460/auue.2025.98.1.332

sam : « jamais » ; du hausa *sám* « pas du tout », via ffde *sam* « particule de renforcement de la négation » (cf. Noye 1989: 299).

tirmi : « pièce de tissu » ; du hausa *túrmíi* « pièce de tissu imprimé de 6 yards », via ffde *turmi(wol)* « coupon de tissu ».

tumtum : « toujours, chaque fois » ; du hausa *túm* « complètement » ; via ffde *tum* « toujours », *tum-tum* « absolument toujours ».

wayna : « beignet plat » ; du hausa *wàináa* « beignet de farine » > kanuri *wainá* « beignet de farine » et ffde *wayna* « beignet rond et plat fait avec du riz ».

5.6 Emprunts du giziga au kanuri via le fulfulde

Le fulfulde du Diamaré a emprunté assez massivement à la langue kanuri (Mohammadou Eldridge 1997). Cela s'explique par le fait que les Peuls arrivant de l'Ouest ont longuement fait escale au Borno, où ils ont emprunté à la fois des mots et des techniques. C'est, ensuite, principalement via le fulfulde que le giziga a emprunté du lexique aux Kanuri. Notre ouvrage de référence pour le kanuri est Cyffer & Hutchison (1990).

Nous remarquons, logiquement, que le giziga a emprunté (indirectement) au kanuri plusieurs termes relatifs à la culture du sorgho repiqué *(bongoro, burgu, hokolori, mosoko)* culture qui a été importée au Diamaré depuis le Borno dans la deuxième partie du 19e siècle (Tourneux & Seignobos 1997: 200 ; Seignobos 2000a, planche 14). De la même façon, plusieurs termes relatifs à l'âne ont été pris par le giziga chez les Kanuri, toujours via le fulfulde *(kalire, kayawa, kilatur)*.

aggalata : « interjection émise pour prendre de l'élan avant un effort particulier », du kanuri *agəlaptə́* « victoire », via ffde *agalata* ! « formidable ! ».

bongoro : « machette pour désherber » ; du kanuri *bongoró* ; cf. ffde *bonngorooru* « couperet ».

burgu : « variété de *muskuwaari* à grain grisâtre » ; du kanuri *búrwu* (**búrgu*) ; > ffde *burguuri* « variété de sorgho repiqué qui donne une farine colorée ».

daldal : « terrain plat » ; du kanuri *dándal* « espace dégagé devant la résidence du *shehu* ou d'un chef », via ffde *daldal* « terrain maintenu sans végétation ; cour ».

galáji : « pois de terre » ; du kanuri *ngángálá* « pois de terre », via ffde *ngalaawu / galaaji* « pois de terre ».

hamham : « sauce faite avec du tourteau d'arachide et de l'oseille de Guinée » ; du kanuri *hám* « tissu beige de qualité inférieure » (Cyffer

DOI 10.15460/auue.2025.98.1.332

& Hutchison 1990: 65) – c'est effectivement la couleur de la sauce dite *hamham* en ffde ; l'étymon kanuri est donc passé par ffde *hamham* « boulette de pâte d'arachide dégraissée cuite à l'eau avec de l'oseille de Guinée » [c'est considéré comme le degré zéro de la cuisine].

hokolori ~ akolori ~ okolori ~ kolori ~ kolokolori : « petite mare artificielle dans un champ à sorgho repiqué » ; du kanuri *kulúwu* « mare, étang », via ffde *okoloore* « petite mare artificielle creusée sur les terrains à sorgho repiqué » [Notons que la technique culturale du sorgho repiqué a été introduite dans la région par les Kanuri (Seignobos 2000, planche 14 « Sorghos et civilisations agraires », dans Seignobos & Iyébi-Mandjek [éd.]).

kal : « peu importe, c'est égal » ; du kanuri *kál* (adj.) « même, similaire, égal » et kanuri *kálkál* « similitude, ressemblance », via ffde *kal* « ça n'a pas d'importance, c'est égal ».

kalire : « âne sp. » ; du kanuri *kəlî* « blanc (cheval) », via ffde *kiliire* « âne à museau blanc ».

kalkal : « égal » ; du kanuri *kál* (adj.) « même, similaire, égal » et *kálkál* « similitude, ressemblance », via ffde *kalkal* « juste, exactement ; ça n'a pas d'importance, c'est égal ».

kasaryel : « store en fines tiges de roseau » ; du kanuri *kásár* « store en tiges de graminées assemblées que l'on suspend généralement dans l'embrasure d'une porte » ; via ffde *kasaryel* « store en tiges de graminées ».

kayawa : « sac en cuir pour âne » ; du kanuri *kăiwa* « grand sac en cuir pour transporter des choses sur un animal » ; cf. ffde *kaywaawu* « sac en cuir ou en tissu que l'on place sur le flanc de l'âne pour y transporter des céréales, de la farine ».

kilatur : « toux des chevaux et des ânes » ; du kanuri *kəlâtərəm* « toux ou rhume, congestion nasale du cheval » ; via ffde *kilaator* « toux ou éternuement du cheval et de l'âne ».

kilbu ~ kulubu : « natron » ; du kanuri *kálwu* (**kálbu*) « natron, potasse » ; via ffde *kilbu*.

mandaŋ : « sel minéral » ; du kanuri *mándá* « sel » ; via ffde *mannda* « sel »

mosoko (dans **daw mosoko**) : « sorgho repiqué » ; < kanuri [**masakwa*] > *mosuwa*, « sorgho de saison sèche » ; cf. ffde *muskuwaari*.

paturu ~ patu : « chat » ; du kanuri *fatú* « chaton », via ffde *paatuuru* « chat ».

sigini : « indigo » ; du kanuri *ságə́nyí* « poudre bleu foncé tirée de l'indigotier et utilisée en teinturerie ; via ffde *sigini* « indigo » ; cf. ffde *siginiiho* « indigotier ».

sokko : « merci » ; du kanuri *wúshe* / plur. *wusówó* (**wusókó*) « salut !, merci ! », via ffde *usoko*, *useko* « merci (à toi).

DOI 10.15460/auue.2025.98.1.332

wayla : « nord », du kanuri *yalá* « nord », via ffde *woyla* « nord ». Le mot ne se trouve que dans le composé **tumuŋ wayla** « grand mouton des Peuls », litt. « mouton du Nord ». Normalement, en giziga, le nord est dit *ti vay muŋ* « vers le haut »).

5.7 Emprunt du giziga à la lingua franca méditerranéenne via le fulfulde

budungaru ~ bidingaru : « fusil » ; de la lingua franca *venediga* « vénitien » ; > arabe *bunduqīya* ; > kanuri *bə́ndə́gə ~ búndúwu* « fusil » > ffde *bunndugaaru* « fusil ». [Jan Knappert (1972: 293), écrit ceci : « La forme hausa *bindiga* vient d'un terme de la *Lingua franca* désignant le fusil *venediga* ; ce mot, en portugais ou en espagnol, est la forme adjectivale de Venise, ancien port de transbordement d'armes à feu ».

5.8 Emprunt du giziga au sango via le fulfulde

makala : « beignet » ; du sango *màkàlà* « beignet de riz, de maïs, de mil ou de blé » (Bouquiaux et collab. 1978: 205) ; via ffde *makala* « beignet de froment » (sans rapport, apparemment, avec l'arabe *akal* « manger ».

5.9 Emprunt du giziga au seereer via le fulfulde

yoloɓiri : « mil très blanc qui se récolte en novembre / décembre » ; du seereer *yoloɓ* « être lâche » (Merrill 2018), via ffde *yoloɓri* « sorgho pluvial tardif à panicules lâches (ou aérées) » (voir Seignobos 2000: 82–87).

5.10 Emprunt du giziga au soŋay via le fulfulde

gaw : « chasseur professionnel » ; du soŋay *Gaw* « nom d'un groupe de chasseurs professionnels » ; via le kanuri *gâu* « médecin local qui vend des plantes et des bois médicinaux », et le ffde *gaw* « chasseur professionnel ». Pour comprendre le passage de « chasseur » à « médecin », voir Seignobos (2011).

5.11 Emprunts du giziga au soninke via le fulfulde

Les références pour cette section sont Diagana (2011) et Tourneux (2014).

bonoy : « douleur, peine, souffrance » ; du soninke *bòné* « malheur », via ffde *bone* « détresse, malheur, souffrance » ou kanuri *báne, bóne* « souffrance, ennuis ».

daago ~ dago : « natte tressée avec des folioles de palmier » ; du soninke *dàagó* « natte en tiges de mil », via ffde du Mali et du Diamaré *daago* « natte tressée en folioles de rônier ou de doum ».

DOI 10.15460/auue.2025.98.1.332

gegeru : « vièle » ; du verbe soninke *gòogó* « gratter » ; via hausa *gòogée* « vièle monocorde », cf. ffde du Mali *googeeru* « vièle monocorde » ; ffde du Diamaré *geegeeru* « vièle monocorde » ; d'après Erlmann (1983: 19), le mot peul est un « étrange mélange d'éléments arabes, hausa et peuls » ; nous n'avons cependant pas pu retrouver, pour l'instant, l'élément arabe évoqué [ġāġā].

goro : « kola » ; < soninke *gòró* (cf. bambara *woro*) ; via ffde du Mali *gooro* et ffde *gooro* « noix de kola » ; cf. hausa *góorò*.

kafaye : « sabre » ; du soninke *káafà* « sabre » via ffde *kaafaahi* « épée, sabre ».

kiyta ~ kiita : « jugement » ; du soninke *kíitè* « procès » ; via ffde du Mali *kiitagol* « jugement ; procès » et ffde *kiita* « jugement ».

5.12 Emprunt du giziga au yoruba via le fulfulde

Notre ouvrage de référence pour le lexique yoruba est Sachnine (2009).

akoti : « caisse en bois ou en métal » ; de façon convaincante, Skinner (1996: 4) met ce mot en relation avec le yoruba *àpótí*, qui désigne, selon Sachnine (2009: 60–61) « tout contenant tel que boîte, caisse, coffre, malle, valise » ; le mot est passé par le hausa *àkwàatì* « boîte, malle, caisse », puis par ffde *akooti* « boîte, caisse, cantine, coffre » avant d'arriver en giziga.

asaana : « allumette » ; du yoruba *àṣánà*, via hausa *àsháanáa* « allumette », puis via ffde *asaana*.

kéeké : « machine à coudre » ; du yoruba *kẹ̀kẹ́* « bicyclette », via hausa *kèekè* « bicyclette ; machine » ; via ffde *keeke* « machine à coudre » ; cf. kanuri *keké* « bicyclette ; machine ».

rekke : « canne à sucre » ; du yoruba *ìrèké* « canne à sucre », via hausa *ràkée* « canne à sucre », via ffde *reke* « canne à sucre ».

5.13 Emprunt du giziga à des langues non déterminées via le fulfulde

Il existe une quinzaine de mots que le giziga a bien empruntés au fulfulde (via ffde), mais dont le fulfulde n'est sans doute pas l'origine ultime.

karanga : « ancienne monnaie » ; via ffde *karanngaare* « ancienne pièce de monnaie (0,25 fr.) ».

kata : « tamis » ; via ffde *kata* « tamis, filtre » (ce mot peul est pour l'instant d'origine inconnue).

DOI 10.15460/auue.2025.98.1.332

kombowal : « pirogue » ; via ffde *koombowal* « pirogue » ; ce mot n'appartient pas au fonds lexical pan-peul ; est donc susceptible d'être un emprunt.

kudaaku : « patate douce » ; via ffde *kudaku* « patate douce » ; probablement d'une langue du sud du Nigeria [cf. yoruba *ọ̀dùnkún* (Sachnine 2009), igbo *kukuǹdùku* (Ugochukwu & Okafor 2004)] via hausa de l'Ader *kúudakúu*.

kundurku ~ hundurku : « boisson à base de farine de mil, légèrement alcoolisée » ; via ffde *kundurku* « boisson légèrement fermentée, un peu acide et épicée ; ce mot est manifestement un emprunt en ffde.

mbay : « manioc » ; via ffde *mbay* « manioc ». D'après Barreteau (1988), le mofu-Gudur *mbáy* « manioc » viendrait du fulfulde. Les récits anciens de voyageurs font état de plantations de manioc aussi bien à Yola qu'à proximité du Logone. D'après Blench (1998, 2014) le manioc, sous le nom de *mbay*, aurait été diffusé dans la région par les Peuls. Il aurait pénétré en Centrafrique à partir de l'Ouest (Cloarec-Heiss & Nougayrol 1998). Il reste que ce mot n'appartient pas au fonds lexical peul. Il faut bien noter, comme nous le rappelait C. Seignobos (c. p. 06.10.2021) que le manioc dont il est question ici est le manioc doux, dont le tubercule se consomme directement, sans rouissage.

ndaga ~ ndagga ~ ndahge ~ ndanga : « variole » ; via ffde *ndagga* « variole », qui est peut-être une innovation locale.

ndifri : « aluminium, étain, plomb » ; via ffde *ndifri* « plomb » ; ce mot n'appartient pas au fonds pan-peul.

ngadiga : « varicelle » ; via ffde *ngaadiga* « varicelle » ; cf. mofu-Gudur *ngáadəga* « varicelle », donné par Barreteau (1988) comme emprunté au ffde.

ngalum : « hameçon » ; via ffde *ngalum* « hameçon », qui l'a emprunté à une langue tierce ; Barreteau (1988) donne *ngálam* « hameçon » comme un emprunt au ffde ; cf. bambara *nɛ̀gɛ̀len* « gros hameçon », *gàngali* « ligne dormante ».

pataka : « ancienne monnaie » ; via hausa *fàtákàa* « pièce de 2 shillings (ancienne monnaie du Nigeria) ; via ffde *pataka* (traduit par « florin » dans Taylor 1932: 154) ; cf. manjaku de Bassarel *patakan* « pièce de monnaie » (Segerer & Flavier 2011–2019). Un expert anonyme nous confirme que de nombreuses langues ouest-africaines disposent du même emprunt : fon (Delafosse 1894: 197) *kpatagã* « argent », vai (Koelle 1854: 206) *pátāwa, pátāra, pátā* « monnaie métallique », emprunt attribuable très probablement au portugais *prata* « argent ».

toro : « ancienne monnaie » ; via ffde *toro* « petite pièce de monnaie (3 pence ? » (mot d'origine inconnue ; donné par R. Blench comme

DOI 10.15460/auue.2025.98.1.332

hausa, dans *A Dictionary of the Basa Language, Trial edition* [Segerer et Flavier 2011–2019]).

zawka : « vaseline » ; via ffde *zawka* « état reluisant du corps humain (en bonne santé) (Noye 1989: 424 » (mot d'origine inconnue).

6 Classement thématique des emprunts au fulfulde ou passés par le fulfulde

Nous allons essayer de classer par thèmes, plus ou moins arbitrairement établis, le vocabulaire non peul que le giziga a emprunté par l'intermédiaire du fulfulde.

6.1 Végétaux

burgu : variété de *muskuwaari* à grain grisâtre
dankali : patate douce
galáji : pois de terre
goro : kola
kudaaku : patate douce
lemu : citron
mbay : manioc
mosoko : sorgho repiqué
rekke : canne à sucre
yoloɓiri : mil très blanc qui se récolte en novembre-décembre

6.2 Animaux

kalire : âne sp.
paturu ~ patu : chat

6.3 Culture matérielle

akoti : caisse en bois ou en métal
asaana : allumette
asusu : caisse à argent
bongoro : machette pour désherber
borgo : couverture
budungaru ~ bidingaru : fusil
daago ~ dago : natte tressée avec des folioles de palmier
danko : fronde en caoutchouc
garaya : "guitare", luth
gegeru : vièle
godo : couverture blanche en coton tissée à la main
goŋgoŋ : boîte ou fût métallique
jawjaw : tambour d'aisselle
jawleeru : vestibule, porte d'entrée
kafaye : sabre
kalanjir ~ kalanjur : lampe en terre fonctionnant avec de la résine
kasaryel : store en fines tiges de roseau
kata : tamis
kayawa : sac en cuir pour âne
kéeké : machine à coudre
kombowal : pirogue
kotóróko ~ kotórko : radier, pont
lammba : numéro ; marque
laya : amulette, grigri
mekef : ciseaux
ndifri : aluminium, étain, plomb
ngalum : hameçon

DOI 10.15460/auue.2025.98.1.332

pitirla : lampe à pétrole
senge : moustiquaire
sigini : indigo
tambal : grand tambour que l'on trouve chez les lamibé
tastilam : lampe torche
urdi : parfum
zawka : vaseline

6.4 Habillement

abada : pagne de piètre qualité
adiku : foulard, mouchoir de tête
alawaya : tunique d'homme à manches courtes
dawra : gandoura
gajeré : culotte
jiyba : poche
sirla : pantalon
tirmi : pièce de tissu

6.5 Alimentation, boisson

arge ~ **erge** : alcool
barasa : alcool, eau de vie
hamham : sauce faite avec du tourteau d'arachide et de l'oseille de Guinée
is : levure
kilbu ~ **kulubu** : natron
kose : sorte de beignet
kundurku ~ **hundurku** : boisson légèrement alcoolisée
kunu : boisson légèrement alcoolisée
makala : beignet
mandaŋ : sel minéral
sa'e : 1. café ou thé ; 2. lie de bière fermentée
sukwar : sucre
wayna : beignet plat

6.6 Agriculture, élevage

daldal : terrain plat
hokolori ~ **akolori** ~ **okolori** ~ **kolori** ~ **kolokolori** : petite mare artificielle dans un champ à sorgho repiqué
kilatur : toux des chevaux et des ânes

6.7 Pratiques et institutions sociales

akre : rétribution en nature ; location
bariki : bureau, administration
dangay : prison
garama : impôt, taxe
gaw : chasseur professionnel
gidaŋmucu : cimetière
kiyta ~ **kiita** : jugement
sariya : jugement
sedege : aumône faite en sacrifice
sedewo : témoin
soje : soldat, gendarme, policier
zakka : dîme, redevance en nature versée aux chefs

DOI 10.15460/auue.2025.98.1.332

6.8 Monnaie, commerce

dala : argent ; unité monétaire valant 5 francs
karanga : ancienne monnaie
koɓo : petite pièce de monnaie
nafa : utilité, intérêt
pataka : ancienne monnaie
riba : gain, bénéfice
sisi : ancienne monnaie
suloy : ancienne monnaie
toro : ancienne monnaie

6.9 Maladies

ndaga ~ ndagga ~ ndahge ~ ndanga : variole
ngadiga : varicelle

6.10 Adverbes, exclamations, éléments grammaticaux

abada : toujours
aggalata : interjection émise pour prendre de l'élan avant un effort particulier
amma : mais
circir très droit
dayday ~ deydey : moyennement ; de taille moyenne
diga : depuis
fakat : exactement, sûrement, certainement
jam : exclamation de contentement, d'agrément
jamjam : en paix, en bonne santé
kal : peu importe, c'est égal
kalkal : égal
kay ! kayya ! : interjection marquant l'étonnement, la surprise, le mécontentement
sam : jamais
sokko : merci
tumtum : toujours, chaque fois
yáwwa(a) : exclamation de contentement
zaaman : autrefois

6.11 Autres lexèmes

asar : malheur, accident, situation triste
bonoy : douleur, peine, souffrance
dabaray : moyen, ruse, stratégie, tromperie
dakare : qui a une mauvaise conduite ; individu qui a une mauvaise conduite
dole : obligatoire, obligé
dulniya ~ duniya : 1. monde ; 2. la vie, les gens
hàràm : interdit, refus catégorique
iriŋ : qualité ; semblable
kono ~ kona ~ kwana : virage, tournant
labara : nouvelle
wayla : nord

DOI 10.15460/auue.2025.98.1.332

7 Tableaux récapitulatifs

Tableau 5. Emprunts du fulfulde au giziga	28
végétaux	9
animaux	9
culture matérielle	4
structure socio-politique	4
autres	2
Tableau 6. Emprunts du giziga au fulfulde	53
végétaux	2
animaux	6
culture matérielle	15
agriculture et élevage	6
pratiques et institutions sociales	13
lexèmes relatifs à la perception	3
adverbes et éléments grammaticaux	8
Tableau 7. Langues des emprunts du giziga via le fulfulde	109
allemand	1
anglais	10
arabe	36
bambara	2
hausa	25
kanuri	20
lingua franca	1
sango	1
seereer	1
soŋay	1
soninke	6
yoruba	4
autres	13

Selon ce décompte et en fonction de notre documentation, le giziga a prêté 28 mots au fulfulde et en a reçu 161 de lui ou par son intermédiaire. Le domaine religieux de l'islam n'ayant pas été exploré, il

DOI 10.15460/auue.2025.98.1.332

est sûr qu'il apporterait un lot supplémentaire d'emprunts à l'arabe via le fulfulde.

8 Liste alphabétique des mots giziga cités dans l'article

abada : toujours
abada : pagne de piètre qualité
adiku : foulard, mouchoir de tête
aggalata : interjection émise pour prendre de l'élan avant un effort particulier
akoti : caisse en bois ou en métal
akre : rétribution en nature ; location
alawaya : tunique d'homme à manches courtes, descendant jusqu'aux genoux
amma : mais
ardedel : *Merremia emarginata* (Convolvulaceae)
arge ~ erge : alcool
arngawo ~ (h)irngawo : lit
asaana : allumette
asar : malheur, accident, situation triste
asusu : caisse à argent
awra : âne sp.
baba : père
bambaɗo : griot
barasa : alcool, eau de vie
bariki : bureau, administration
batal : aiguille ; seringue ; injection
bo : aussi, également
bodogor : célibataire
bohal ~ bohol ~ bowal ~ buwal ~ bowagol : grand-route
bongoro : machette pour désherber
bonoy : douleur, peine, souffrance
borgo : couverture
bortol : chemin par où passe le bétail
budungaru ~ bidingaru : fusil
buɗare : vagabond
bugawla : [litt : chef jeune homme], vainqueur, le plus fort ; héros
burgu : variété de *muskuwaari* à grain grisâtre
ɓamtal : poteau central
ɓesde : ajout, cadeau qu'on ajoute sur ce qu'on a acheté
ɓokki : baobab
caca ~ caaca : jeu de hasard
caɗ(u)ngu : animal mort sans avoir été égorgé
caka cak : juste par le milieu, juste au milieu
celelew : chaîne en fer
cete : brochette de viande
circir très droit
daago ~ dago : natte tressée avec des folioles de palmier
dabaray : moyen, ruse, stratégie, tromperie
dakare : qui a une mauvaise conduite ; individu qui a une mauvaise conduite
dala : argent ; unité monétaire valant 5 francs
daldal : terrain plat
dangay : prison
dankali : patate douce
danko : fronde en caoutchouc
darogal ~ dorogal : miroir, glace
dawra : gandoura
dayday ~ deydey : moyennement ; de taille moyenne
diga : depuis
dimu : liberté
dole : obligatoire, obligé

DOI 10.15460/auue.2025.98.1.332

dulniya ~ duniya : 1. monde ; 2. la vie, les gens
dus : odeur nauséabonde
ɗerewol ~ ɗerewel : papier ; livre, cahier
fakat : exactement, sûrement, certainement
galáji : pois de terre
garama : impôt, taxe
garaya : “guitare”, luth
gaw : chasseur professionnel
gáwla : jeune homme
gegeru : vièle
gajeré : culotte
gidaŋmucu : cimetière
giɗaɗo : épouse courageuse au travail
deno-deno : bleu
godo : couverture blanche en coton tissée à la main
goŋgoŋ : boîte ou fût métallique
goro : kola
gudi : forgeron
gula : gauche ; gaucher
gulek : pioche
hamham : sauce faite avec du tourteau d’arachide et de l’oseille de Guinée
hàràm : interdit, refus catégorique
hokolori ~ akolori ~ okolori ~ kolori ~ kolokolori : petite mare artificielle dans un champ à sorgho repiqué
horende : chéchia de couleur rouge
iriŋ : qualité ; semblable
is : levure
jaɓama : bienvenu ; bienvenue !
jam : exclamation de contentement, d’agrément
jamjam : en paix, en bonne santé
jawal : rapidité, précipitation
jawjaw : tambour d’aisselle
jawleeru : vestibule, porte d’entrée
jiyba : poche , de l’arabe *jībat* poche
jiyku : bonnet
jobol : ligne, rang, rangée
kafaye : sabre
kal : peu importe, c’est égal
kalkal : égal
kalanjir ~ kalanjur : lampe en terre fonctionnant avec de la résine
kalire : âne sp.
karal : sol argileux pour sorgho de contre-saison
karanga : ancienne monnaie
karawal : siège, chaise
kasaryel : store en fines tiges de roseau
kata : tamis
kay ! kayya ! : interjection marquant l’étonnement, la surprise, le mécontentement
kayawa : sac en cuir pour âne
kéeké : machine à coudre
kilatur : toux des chevaux et des ânes
kilbu ~ kulubu : natron
kiyta ~ kiita : jugement
koɓo : petite pièce de monnaie
kocori ~ kocoro ~ kocorop ~ hosori ~ husuru : 1. escargot ; 2. coquille d’escargot
kombowal : pirogue
kono ~ kona ~ kwana : virage, tournant
kose : sorte de beignet
kotóróko ~ kotórko : radier, pont
kucum : daman des rochers
kudaaku : patate douce
kumambeɗe : herbe sp.
kundurku ~ hundurku : boisson à base de farine de mil, légèrement alcoolisée

DOI 10.15460/auue.2025.98.1.332

kunu : boisson à base de farine de mil, légèrement alcoolisée
kurum : très noir
labara : nouvelle
lammba : numéro ; marque
laya : amulette, grigri
lemu : citron
lugere : bonne terre, terre fertile
luma ~ lumo : marché ; semaine
maaba : coucal du Sénégal
makala : beignet
malalay : poisson sp. très glissant, avec des nageoires piquantes
maliya : *Ficus thonningii* (Moraceae)
malla : ou bien
mandaŋ : sel minéral
mangaraw : *Corchorus tridens* (Tiliaceae)
marawo : gifle
maray ~ muray poisson à queue à moitié rouge qui ressemble à celle de la “sardine”
mbaɮa : bière de mil
mbay : manioc
mbere-mbere ~ membere-mbere maja : *Commelina benghalensis* (Commelinaceae)
mbirlek ~ mburleke : grenouille
mbolori ~ mbulori : chapeau de paille
meɓɓere : foule
mekef : ciseaux
memel ~ memeɗ *Corchorus fascicularis* (Tiliaceae)
metalel : turban
menjeɗe ~ manjaɗay : petit poisson gras
mepelepele : poisson à chair rouge, gras et plein d’arêtes
mindek : Ficus sp.
mogoyok : cendres de tiges de mil
mohol : mur d’enceinte
molde : ânon
mosoko (dans **daw mosoko**) : sorgho repiqué
murla ~ mərla ~ mirla : colostrum
muzuk : variété de sorgho rouge pluvial
nafa : utilité, intérêt
nara : entente
ndaga ~ ndagga ~ ndahge ~ ndanga : variole
ndifri : aluminium, étain, plomb
ndola : poisson noir allongé [sans écailles]
ngaari : taurillon
ngadiga : varicelle
ngalalay : *Dalbergia melanoxylon*
ngalum : hameçon
ngorgi : même classe d’âge, classe d’initiation
njáɓ : idéophone : qui se conviennent physiquement ou moralement
pataka ancienne monnaie
paturu ~ patu : chat
petengew : petite grenouille à gros ventre
pitirla : lampe à pétrole
rekke : canne à sucre
riba ; gain, bénéfice
saɓɓa : paille, herbe qu’on coupe pour aller la mettre dans le champ de karal [terrain à sorgho repiqué] avant d’y mettre le feu
sa’e : 1. café ou thé ; 2. lie de bière fermentée
sakre : pépinière (en particulier de sorgho de contre-saison)
sam : jamais
sayre : bonnet en toile à matelas
sariya : jugement
sedege : aumône faite en sacrifice

DOI 10.15460/auue.2025.98.1.332

sedewo : témoin
senge moustiquaire
sigini : indigo
siké : alors, vraiment
siriw : silencieux
sirla : pantalon
sisi : ancienne monnaie
soje : soldat, gendarme, policier
sokko : merci
sukwar (dans *mandaŋ sukwar*) : sucre
suloy : ancienne monnaie
ta(a)ri : fronde en caoutchouc
tal tal : tout blanc (d'une certaine étendue)
tambal : grand tambour que l'on trouve chez les lamibé
tirmi : pièce de tissu
toro : ancienne monnaie
tastilam : lampe torche
tumtum : toujours, chaque fois
turus : *Haematostaphis barteri*
urdi : parfum
wayla : nord (dans *tumuŋ wayla* grand mouton des Peuls, litt. mouton du Nord)
wayna : beignet plat
yáwwa(a) : exclamation de contentement
yoloɓiri : mil très blanc qui se récolte en novembre / décembre
zaaman : autrefois
zakka : dîme, redevance en nature versée aux chefs
zawka : vaseline
zindirɗa : ver de terre

Bibliographie

Baldi, Sergio. 2008. *Dictionnaire des emprunts arabes dans les langues de l'Afrique de l'Ouest et en swahili*. Paris: Karthala.

Baldi, Sergio. 2015. *Dizionario Hoepli Hausa, Hausa-Italiano, Italiano-Hausa.* Milan: Hoepli.

Bargery, George Percy. 1993 (2e éd.). *A Hausa-English dictionary and English-Hausa vocabulary, with some notes on the Hausa people and their language* by Diedrich Westermann, *and Supplement* by A. Neil Skinner. Zaria: Ahmadu Bello University Press.

Barreteau, Daniel. 1988. *Description du mofu-gudur (langue de la famille tchadique parlée au Cameroun). Livre 2. Lexique.* Paris : Éditions de l'ORSTOM.

Barreteau, Daniel & Michel Dieu. 2000. Linguistique. In Christian Seignobos & Olivier Iyébi-Mandjek (dir.), *Atlas de la province Extrême-Nord Cameroun.* Paris : IRD Éditions ; Yaoundé : MINREST/INC. 64–65.

Barreteau, Daniel & Yves Le Bléis. 1991. *Lexique mafa (langue de la famille tchadique parlée au Cameroun)*. Paris : Geuthner/ORSTOM.

Blache, Jacques, François Miton, Alfred Stauch, André Iltis & Gérard Loubens. 1964. *Les Poissons du bassin du Tchad et du bassin adjacent du Mayo Kebbi : Étude systématique et biologique.* Paris: ORSTOM.

DOI 10.15460/auue.2025.98.1.332

Blench, Roger. 1998. « The introduction and spread of New World crops in Nigeria : a historical and linguistic investigation ». In Monique Chastenet (dir.), *Plantes et paysages d'Afrique : une histoire à explorer*, Paris **:** Karthala. 165–210.

Blench, Roger. 2014. The diffusion of cassava in Africa: lexical and other evidence. Ma.

Bouquiaux, Luc (avec la collaboration de Jean-Marie Kobozo & Marcel Diki-Kidiri). 1978. *Dictionnaire sango-français.* Paris: Société d'Études Linguistiques et Anthropologiques de France (SELAF).

Cloarec-Heiss, France & Pierre Nougayrol. 1998. Des noms et des routes : La diffusion des plantes américaines en Afrique centrale (RCA – Tchad). In Monique Chastenet (dir.), *Plantes et paysages d'Afrique : une histoire à explorer*. Paris **:** Karthala. 117–164.

Cyffer, Norbert & John Hutchison (éd.). 1990. *Dictionary of the Kanuri language*. Dordrecht: Foris; Maiduguri: University of Maiduguri.

Delafosse, Maurice. 1894. *Manuel dahoméen: grammaire – chrestomathie, dictionnaire français-dahoméen et dahoméen-français.* Paris : Ernest Leroux.

De Wolf, Paul Polydor. 1995. *English-Fula dictionary (Fulfulde, Pulaar, Fulani): A multidialectal approach*, 3 vol. Berlin: Dietrich Reimer.

Diagana, Ousmane Moussa. 2011. *Dictionnaire soninké-français*. Paris: Karthala.

Dumestre, Gérard. 2011. *Dictionnaire bambara-français, suivi d'un abrégé français-bambara*. Paris : Karthala.

Erlmann, Veit. 1983. Notes on musical instruments among the Fulani of Diamaré (North Cameroon). *African Music* 6(3). 16–41.

Gouffé, Claude. 1971. Notes de lexicologie et d'étymologie soudanaises : I. A propos du nom du moustique et de la moustiquaire. *Comptes rendus du Groupe linguistique d'études chamito-sémitiques G.L.E.C.S.* 15. 55–65.

Gravina, Richard. 2014. *The phonology of Proto-Central Chadic: the reconstruction of the phonology and lexicon of Proto-Central Chadic, and the linguistic history of the Central Chadic languages.* Utrecht: LOT.

Guitard, Émilie. 2017. Guiziga Bui Marva. In Saheed Aderinto (ed.), *African kingdoms: An encyclopedia of empires and civilizations.* Santa Barbara: ABC-CLIO. 107–110.

Jaggar, Philip J. 2010. The role of comparative/historical linguistics in reconstructing the past: What borrowed and inherited words tell us about the early history of Hausa. In Anne Haour & Benedetta Rossi

DOI 10.15460/auue.2025.98.1.332

(eds.), *Being and becoming Hausa (interdisciplinary perspectives)*. Leiden & Boston: Brill. 35–58.

Jaouen, René. 1990. Exposé sommaire sur les verbes monoradicaux en giziga-sud (giziga de Muturwa, Cameroun). In Herrmann Jungraithmayr & Henry Tourneux (éd.), *Études tchadiques: verbes monoradicaux*. Paris: Geuthner. 215–219.

Jaouen, René. 1995. *L'Eucharistie du mil: langages d'un peuple, expressions de la foi*. Paris: Karthala.

Jullien de Pommerol, Patrice. 1999. *Dictionnaire arabe tchadien – français suivi d'un index français-arabe et d'un index des racines arabes*. Paris: Karthala.

Knappert, Jan. 1972. The study of loan words in African languages, *Afrika und Übersee* 56(4). 283–308.

Koelle, Sigismund Wilhelm. 1854. *Outlines of a grammar of the Vei language, together with a Vei-English vocabulary*. London: Church Missionary House.

Lacroix, Pierre-Francis. 1967. Remarques préliminaires à une étude des emprunts arabes en peul. *Africa* 37(2). 188–202.

Lukas, Johannes. 1970. *Studien zur Sprache der Gisiga (Nord-Kamerun)*. Glückstadt-Hamburg: J.J. Augustin.

Merrill, John. 2018. Sereer-English / English-Sereer dictionary. Berkeley: University of California.

Michielan, Antonio et al. Dictionnaire giziga-français. Ma.

Mohammadou Eldridge (éd.). 1976. *L'Histoire des Peuls Férôbé du Diamaré : Maroua et Petté*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Mohammadou Eldridge. 1997. Kanuri imprint on Adamawa Fulbe and Fulfulde. In Norbert Cyffer & Thomas Geider (eds.), *Advances in Kanuri scholarship*. Köln: Rüdiger Köppe Verlag. 257–311.

Newman, Paul. 2007. *A Hausa-English dictionary*. New Haven & London: Yale University Press.

Newman, Paul & Roxana Ma Newman (compil.). 1982 [1re éd. 1977]. *Modern Hausa-English dictionary / Sabon Kamus Na Hausa Zuwa Turanci*. Ibadan: University Press.

Newman, Paul & Roxana Ma Newman. 2020. *Hausa dictionary for everyday use: Hausa-English/English-Hausa, Kamusun Hausa na Yau da Kullum: Hausa-Ingilishi/Ingilishi-Hausa*. Kano: Bayero University Press.

DOI 10.15460/auue.2025.98.1.332

Noye, Dominique. 1989. *Dictionnaire foulfouldé-français, dialecte peul du Diamaré, Nord-Cameroun*. Préface de R. Labatut, illustrations de C. Seignobos. Garoua: Procure des Missions ; Paris: P. Geuthner.

Parietti, Giuseppe & Henry Tourneux (collab.). 2018. *Dictionnaire fulfulde-français / français-fulfulde* (Dialect[e] peul [du] Diamaré, Cameroun). Illustrations de Christian Seignobos. Pessano con Bornago: Mimep-Docete.

Pontié, Guy. 1973. *Les Guiziga du Cameroun septentrional : L'organisation traditionnelle et sa mise en contestation.* Paris: ORSTOM.

Pontié, Guy. 1981. Quelques éléments d'histoire guiziga. In Claude Tardits (dir.), *Contribution de la recherche ethnologique à l'histoire des civilisations du Cameroun*, vol. 1. Paris: CNRS. 249–263.

Pozdniakov, Konstantin. 2022. *Proto-Fula-Sereer: Lexicon, morphophonology, and noun classes*. Berlin: Language Science Press.

Sachnine, Michka. 2009. *Dictionnaire usuel yorùbá-français, suivi d'un index français- yorùbá*. Paris : Karthala ; Ibadan : IFRA.

Segerer, Guillaume & Sébastien Flavier. 2011–2019. *RefLex: Reference Lexicon of Africa*, Version 1.2. Paris, Lyon. http://reflex.cnrs.fr/

Seghieri, Josiane. Sans date. Flore guiziga. Document inédit.

Seghieri, Josiane. 1990. *Dynamique saisonnière d'une savane soudano-sahélienne au Nord-Cameroun.* Montpellier : Université de Montpellier II, Thèse de doctorat.

Seghieri, Josiane & Christian Floret. 1993. Dynamique saisonnière de la végétation en savane sahélo-soudanienne. In *Cahier scientifique* 11, Les Terres hardé. Nogent-sur-Marne: CIRAD-Forêt. 55–64.

Seignobos, Christian. 2000a. Sorghos et civilisations agraires. In Christian Seignobos & Olivier Iyébi-Mandjek (dir.), *Atlas de la province Extrême-Nord Cameroun*. Paris: IRD Éditions ; Yaoundé: MINREST/ INC. 82–87.

Seignobos, Christian. 2000b. Fulbe. In Christian Seignobos & Olivier Iyébi-Mandjek (dir.), *Atlas de la province Extrême-Nord Cameroun.* Paris: IRD Éditions ; Yaoundé: MINREST/ INC. 52–56.

Seignobos, Christian. 2011. Les gaw du Cameroun, de chasseurs à tradipraticiens. *Journal des africanistes* 81(1). 35–59.

Seignobos, Christian. 2014. La chasse/pêche aux Batraciens : aux origines de la vie des populations du bassin du lac Tchad ? (L'exemple du Diamaré, Cameroun). *Anthropozoologica* 49(2). 305–325.

Seydou, Christiane. Sous presse. *Dictionnaire peul du Mali*.

DOI 10.15460/auue.2025.98.1.332

Shay, Erin. 2021. *A grammar of Giziga: A Chadic language of Far North Cameroon.* (Grammars and Sketches of the World's Languages 14). Leiden: Brill.

Skinner, Neil. 1996. *Hausa comparative dictionary*. Köln: Rüdiger Köppe Verlag.

Taylor, Frank William. 1932. *A Fulani-English dictionary.* Oxford: Clarendon Press.

Tourneux, Henry. 1984. Vocabulaires comparés des instruments aratoires dans le Nord-Cameroun. *Cahiers de l'ORSTOM*, Série Sciences Humaines 20(3–4). 597–612.

Tourneux, Henry. 1991. Observations sur le nom du fer et de la forge dans le bassin du lac Tchad. In Yves Moñino (éd.), *Actes du IVe colloque Méga-Tchad, CNRS/ORSTOM, Paris, du 14 au 16 septembre 1988. Vol. 1. Forge et forgerons*. Paris: ORSTOM. 16–30.

Tourneux, Henry, avec la collaboration de Boubakary Abdoulaye, Hadidja Konaï & Fakih Ousmane. 2007. *Dictionnaire peul du corps et de la santé (Diamaré, Cameroun)*. Paris: OIF/Karthala.

Tourneux, Henry. 2013. Les noms de poissons en kotoko commun. In Henry Tourneux (ed.), *Topics in Chadic Linguistics VII, papers from the 6th Biennial International Colloquium on the Chadic Languages, Villejuif, September 22–23, 2011.* Köln: Rüdiger Köppe Verlag. 185–201.

Tourneux, Henry. 2014. Les vestiges du lexique soninke dans le fulfulde du Diamaré (Cameroun). In Ursula Baumgardt & Abdourahmane Diallo (éd.), *La transmission culturelle : l'exemple du peul.* Paris: Karthala. 69–90.

Tourneux, Henry & Christian Seignobos. 1997. Origine et structure du lexique botanique peul du Diamaré. In Daniel Barreteau, René Dognin & Charlotte von Graffenried (éd.), *L'Homme et le milieu végétal dans le bassin du lac Tchad.* Paris : Éd. de L'ORSTOM. 195–216.

Tourneux, Henry & Yaya Daïrou, avec la collaboration de Boubakary Abdoulaye. 2017. *Dictionnaire peul encyclopédique de la nature (faune/flore), de l'agriculture, de l'élevage et des usages en pharmacopée (Diamaré, Cameroun), suivi d'un index médicinal et d'un index français-fulfulde.* Yaoundé: CERDOTOLA.

Tourneux, Henry, Yaya Daïrou & Mohamadou Ousmanou. 2021 [paru en mars 2022]. *Vocabulaire peul du Cameroun, Maroua – Garoua – Ngaoundéré.* Maroua, Faculté des Arts, Lettres et Sciences humaines.

Ugochukwu, Françoise & Peter Okafor. 2004. *Dictionnaire igbo-français suivi d'un index français-igbo.* Paris & Ibadan: Karthala/IFRA.

DOI 10.15460/auue.2025.98.1.332

Wolff, H. Ekkehard. 2022. *A historical phonology of Central Chadic: Prosodies et lexical reconstruction.* Cambridge: Cambridge University Press.

DOI 10.15460/auue.2025.98.1.332

An overview of relative clauses in Gavar

Melanie Viljoen

La Trobe University & SIL Cameroon
M.Viljoen@latrobe.edu.au

Abstract
The goal of this paper is to give a typological overview of the main characteristics of restrictive relative clauses in Gavar. The analysis is based on 572 examples of relative clauses taken from a corpus of 59 interlinearised texts, elicited examples and example sentence provided for lexical items. It is observed that relative clauses in Gavar are externally headed, with the head noun, when present, preceding the relative clause. Headless relative clauses are also possible. The relative clause begins with a relative marker. Subject relative clauses are marked with the relative marker *ma* whose tone varies between high and low. Non-subject relative clauses are marked with relative marker *ndə́ (ná)* which always carries a high tone. There is no restriction on the grammatical roles that can be relativised. Subject, direct and indirect object, oblique, genitive and object of comparison relative clauses are all possible. The gap strategy is used to encode NP_{rel} within a subject relative clause. A resumptive pronoun is frequently but not always used to encode NP_{rel} within a non-subject relative clause. Gavar does not have special 'relative' tense/aspect marking, but perfective marking on the verb is disallowed in relative clauses. The existential status of the head noun is coded by the use versus the absence of a post-relative demonstrative. Subject relative clauses with an adjectival predicate form the basis of the comparative construction in Gavar. Non-subject relative clause constructions have developed into various types of adverbial clauses.

Keywords: relative clause, Chadic language, adverbial clause, comparative construction

1 Introduction

This introduction contains a brief outline of the background of the Gavar language and previous linguistic research done on the language (Section 1.1). The methodology used is described in Section 1.2, and Section 1.3 provides an outline of the rest of the paper.

DOI 10.15460/auue.2025.98.1.354

1.1 Language background

Gavar is a Central Chadic language spoken by an estimated population of 15,000. The Gavar homeland is located in the subdivision of Mogode, division of Mayo-Tsanaga, in the Far North Region of Cameroon. Sizable populations of ethnic Gavar can be found in the Hina subdivision to the south, Garoua (the capital of the North Region), and across the border in Nigeria.

The classification of this language as reported in the *Ethnologue* (Eberhard, Simons & Fennig 2024) is Afro-Asiatic, Chadic, Biu-Mandara, A, A.7. According to Gravina's (2014: 26) revised internal classification of Chadic Biu-Mandara, Gavar belongs to the South subbranch, the Daba Group and the Buwal subgroup. The Biu-Mandara branch of Chadic is referred to by some authors as Central (Gravina 2014: 8).

A limited number of linguistic studies covering the phonology and some aspects of the grammar have been done on the Gavar language (Noukeu 2004; Tchikoua 2006; Viljoen 2017a; Viljoen 2019; Viljoen 2022).

1.2 Methodology

The current study was undertaken based on data collected in the Gavar language area and with a speaker in the capital city of Yaounde, in the period of 2006–2018, by both the author and a colleague, Jolanda Walhout, as part of their work with the Gavar community under the auspices of SIL Cameroon, a branch of SIL International. They were privately funded and have no conflicts of interest. The textual data was provided by numerous members of the Gavar community. Particular acknowledgment, however, goes to Ngama Paul and Kodji Emmanuel of Kwitakwa and Kodji Glab Athanas of Melehleh who also assisted with the transcription and annotation of the textual data as well as providing much of the elicited data.

The analysis of relative clauses is based on 572 examples found in a corpus created by the author, containing the data collected as described above. The corpus consists of 59 interlinearised texts of various lengths and genres (narrative, expository, procedural and hortatory), numerous elicited examples, and illustrative example sentences provided for lexical items. All of this data has been transcribed and annotated in FieldWorks Language Explorer (FLEx) (database not

DOI 10.15460/auue.2025.98.1.354

yet publicly available). Language examples throughout this paper are transcribed phonemically. Tone is not marked on the verb as verbs in Gavar are underlyingly toneless (Viljoen 2019: 258). Tone marking is also omitted from certain borrowings from other languages such as French proper nouns.

1.3 Outline of paper

This paper provides a descriptive overview of restrictive relative clauses in Gavar. Section 2 gives a brief morphosyntactic profile of Gavar which will assist in interpreting the Gavar language examples provided. Section 3 examines how relative clauses can be categorised according to various typological parameters. Section 4 outlines the various types of adverbial clauses in Gavar which have developed from the relative clause structure. Comparative clauses also based on the relative clause structure are described in Section 5. A conclusion will be given in Section 6.

2 Morphosyntactic profile of Gavar

The basic constituent order of Gavar active verbal clauses is SVO/SV. This order may vary under certain pragmatic conditions. Copula clauses may either be verbal or non-verbal and have the following structure: subject (COP) predicate. Copula clauses with no copula are used with equational predicate nominals and predicate adjectives. The focus copula *ànà* is restricted to equational predicate nominals. It shifts the focus to the subject of the copula clause and forms the basis of pseudo-cleft constructions (Viljoen 2017a: 155–159). The copulative verb *da* ‘to be’ is used with non-equational predicate nominals such as proper inclusion, characterisation and specification. It can also be used with predicate adjectives and prepositional phrases expressing location and possession. The verb *ndza* ‘to sit/stay’ can also function as a copula for locative and possessive copula clauses.

Nominal morphology is limited to a plural suffix on a small set of animate nouns and genitive marking on kinship nouns. Verb morphology is far more complex and includes person/number marking, tense-aspect-mode (TAM) marking and other derivational extensions (Viljoen 2019). Like other languages in the area, Gavar has sentence-final negation particles and interrogative words. There is also a negative auxiliary *nkʷɔ́lá* which may precede the main verb. Bipartite

DOI 10.15460/auue.2025.98.1.354

negation involving both the negative auxiliary and a sentence-final negation particle is possible, but rare. Gavar is a tonal language, with high- and low-level tones but as for many Chadic languages, the lexical load is small. Verbs are underlyingly toneless. There is some grammatical variation in tone and any such tonal changes related to relative clauses will be described below.

3 The typology of Gavar relative clauses

The analysis of relative clauses in Gavar is based on a number of typological parameters which are outlined below.

A relative clause can be simply defined as "[a] type of clause, most often a subordinate clause, which serves to modify a noun phrase" (Trask 1993: 238). A more detailed definition is given by Andrews (2007: 206): "A relative clause is a subordinate clause which delimits the reference of an NP by specifying the role of the referent of that NP in the situation described by the RC.[1]"

Andrew's definition applies in particular to restrictive relative clauses which identify the referent of an NP by restricting the possible set of referents to a subset that the speaker wants to talk about (Trask 1993: 238; Tallerman 2015: 276). It is also possible to have non-restrictive relative clauses which "merely adds further information about the NP, without being required for identification" (Trask 1993: 238).

The vast majority of relative clauses found in the Gavar corpus are restrictive and therefore the analysis presented here will concern these. Only one example of a possible non-restrictive relative clause was found in the Gavar corpus (1). This has the same structure as restrictive relative clauses, although there appears to be a pause after the head noun which is not present in restrictive relative clauses.

(1) *gʷàjgʷàjà* *gəwlà,* *mà* *ká-pew* *á* *xàjàk*
festival Gula REL.SBJ IPFV-happen LOC land
gàvàr
Gavar
'…the festival of Gula, which is happening in the land of Gavar'

1 NP = noun phrase, RC = relative clause

DOI 10.15460/auue.2025.98.1.354

Relative clauses may or may not be embedded in the main clause. When they are embedded, the relative clause and the head noun form a complex NP that then occupies a standard NP position in the main clause (Tallerman 2015: 280). Non-embedded relative clauses occur separately from the head noun and may be adjoined to the left or right of the main clause (Andrews 2007: 2014–2015). All the examples in the Gavar corpus are of embedded relative clauses. At this point, it is unknown whether adjoined relative clauses are possible.

Relative clauses have the following parts (Payne 1997: 325–326):

1. The head noun – the NP modified by the relative clause
2. The restricting clause – the relative clause itself
3. The relativised noun phrase (NP_{rel}) – the element within the relative clause that is coreferential with the head noun
4. The relative marker – the morpheme or particle that marks the restricting clause as being a relative clause

Note that a relative pronoun is a relative marker that reflects properties of NP_{rel}, for example its grammatical role.

Relative clauses can be categorised according to a number of typological parameters (Payne 1997: 326; Andrews 2007: 207; Tallerman 2015: 279–283):

1. The order of the relative clause and the head noun.
2. The presence or absence of a relative marker.
3. The strategy used to express NP_{rel} within the relative clause.
4. Which grammatical relations can be relativised.

With regard to the grammatical relations which can be relativised in a language, reference is often made in the literature to “the accessibility hierarchy”, an example of which is given in Fig. 1 below. The idea is that in a particular language, if the NP_{rel} can bear one of the functions in the hierarchy, it can also bear all the functions to the left.

Subject > Direct Object > Indirect Object > Oblique > Genitive > Object of Comparison

Figure 1: Accessibility hierarchy of grammatical relations of NP_{rel} (Keenan & Comrie 1977: 66)

Another important typological parameter for the analysis of relative clauses, not mentioned above, is tense/aspect marking. Chadic languages frequently have verb forms in relative clauses which are different from verb forms in indicative simple sentences. These ‘relative’

DOI 10.15460/auue.2025.98.1.354

forms may also appear in other types of embedded clauses (Frajzyngier 1996: 454–455; Zima 1991).

Frajzyngier (1996: 421–422) also notes that Chadic languages generally encode what he calls the 'existential status' of the head noun, meaning that its existence either has or had not been previously established. This roughly corresponds with the definiteness/indefiniteness distinction. He hypothesises that this distinction in Chadic languages is encoded by one of three possibilities:

1. Contrast between two relative markers
2. The use versus the absence of post-relative markers
3. The use versus the non-use of a relative marker

Other structures that often resemble relative clauses cross-linguistically are questions, adverbial clauses and comparatives (Thompson et al 2007: 246–247, 249–250; Andrews 2007: 207). In Gavar relative clauses form the basis of certain adverbal clauses as well as comparative constructions.

What follows is a categorisation of Gavar relative clauses according to typological parameters outlined above. The order of the relative clause and the head noun is described in Section 3.1. The nature of the relative marker is explored in Section 3.2. Section 3.3 addresses the strategies used to express the relativised noun phrase within the relative clause while Section 3.4 outlines which grammatical relations can be relativised. The tense/aspect marking of verbs within relative clauses is addressed in Section 3.5. Section 3.6 examines the question of how Gavar encodes the existential status of the head noun.

3.1 The order of the relative clause and the head noun

Like most VO languages (Payne 1997: 326), all Gavar relative clauses follow the head noun (HN), when present, as seen in examples (2a) and (2b) below. The head noun is external to the relative clause.

(2)a.	[*skə̀n*]HN	[*mà*	*mbla-xà*	*á*	*làlə̀x*^{w}]RC	*tá-mtʃa*
	thing	REL.SBJ	catch-VNT	LOC	hunt	PFV-die

DOI 10.15460/auue.2025.98.1.354

tə̀.[2]
3SG.POSS
'The animal which was caught at the hunt died.'

b. [*rə̀kʷə̀ts*]HN [*ndə́ sá-tsa-ə̀*]RC *wàtsə́*
garment REL.NSBJ 1SG.SBJ-put-3SG.DOBJ DEM.PROX
tʃáp ábə̀ rà.
cover.entirely with arm
'This garment that I put on has long sleeves.'

The same order of head followed by relative clause is found in nearby Chadic languages such as Buwal (Viljoen 2013: 553–554), Mina (Frajzyngier & Johnston 2005: 429), Daba (Lienhard & Giger 1976: 15), Mofu-Gudur (Hollingsworth 2004: 42) and Cuvok (Dadak 2021: 181).

Headless relative clauses (called 'free' by Andrews (2007: 208)) are also common in Gavar. In such clauses, the head noun is absent and the relative clause refers directly to the referent in question. In Gavar, headless relative clauses can occur in either the subject (3a) or the predicate position (3b) in copula clauses. Note that (3b) is an example of a pseudo-cleft construction based on a copula clause containing the focus copula *ànà* as described in Section 2.

(3)a. [*mà ɓin-xà àká à:gí*]RC *blə̀kʷ wàm*
REL.SBJ be.left-VNT IMM PL thousand ten
áxà nfáɗ.
over four
'There were fourteen thousand left.' (lit. 'Those who were left were fourteen thousand.')

b. *ŋgàlèlèm ànà* [*ndə́ wàlà*
gravel COP REL.NSBJ woman

2 As is common in many Chadic languages, for some intransitive clauses the verb is followed by an intransitive copy pronoun (ICP) which agrees with the subject, and in the case of Gavar, has the same form as the possessive adjective (Viljoen 2017b: 52–54).

DOI 10.15460/auue.2025.98.1.354

à-ɮaɗ-à-zə́ *á* *dʒàk*.]RC
3SG.SBJ-sweep-VNT-TRANS LOC hut
'It is gravel that the woman sweeps from the hut.' (lit. 'Gravel is that which the woman sweeps from the hut.')

Headless relative clauses are also possible in neighbouring languages (Viljoen 2013: 553–554; Frajzyngier & Johnston 2005: 431; Lienhard & Giger 1976: 15).

Since Gavar has sentence-final question words, the headless subject relative clause forms the basis of questions about the subject as in examples (4a) and (4b). In this structure, the relative clause becomes the subject of a copula clause, and the interrogative word the predicate. This same structure can be observed in the closely related language of Buwal (Viljoen 2013: 504–505).

(4)a. [*mà* *vəla-zə́-àkà* *gʷàblà* *tə́*]RC *vàjè*?
REL.SBJ give-TRANS-1SG.IOBJ hide 3SG.POSS who
'Who (will) give me its hide?' (lit. 'The one who (will) give me its hide is who?')

b. [*mà* *ɗal-àtànà*]RC *vémé*?
REL.SBJ do-3PL.IOBJ what
'What happened to them?' (lit. 'That which happened to them is what?')

In active verbal clauses, headless relative clauses can function as subject (5a), direct object (5b) and indirect object (5c).

(5)a. [*mà* *ká-ʔʷən* *dʒàk,*]RC *tà-hà.*
REL.SBJ IPFV-build house arrive-VNT
'The one who is building the house has arrived.'

b. *à-gam* [*má* *ká-ɗal* *pìw.*]RC
3SG.SBJ-drive.away REL.SBJ IPFV-do disorder
'He drives away the one who is creating disorder.'

c. *sà-tá-vəla-àtànà* *dàlà* *á* [*má* *ká-ɗal*
1SG.SBJ-PFV-give-3PL.IOBJ money LOC REL.SBJ IPFV-do

DOI 10.15460/auue.2025.98.1.354

ɮə̀n *à:gí.*]RC
work PL
'I gave money to those doing work.'

3.2 The relative marker in Gavar

Relative clauses in Gavar are marked with a relative marker which follows the head noun (where present) and begins the relative clause. In fact, the relative marker could be better termed a relative pronoun since it carries some information about the grammatical role of the NP within the relative clause. In this regard, Gavar distinguishes between subject and non-subject. A distinction between subject and non-subject in Gavar can also be seen in terms of person/number markers attached to the verb, subject markers being prefixed and other grammatical roles (direct object, indirect object, comitative) being suffixed to the verb (Viljoen 2019: 260–263).

It is not unusual for Central Chadic languages to encode the role of the head noun through a distinction in the form of the relative marker (Frajzyngier 1996: 437). In fact, a distinction between subject and non-subject is also seen in nearby languages such as Daba (Lienhard & Giger 1976: 15), Cuvok (Dadak 2021: 179) and Mofu-Gudur (Hollingsworth 2004: 42–44).

The subject relative marker in Gavar has the form *ma*. This is likely to have developed from the common *m-* nominalising prefix found in Chadic and in Afro-asiatic languages in general (Greenberg 1966: 48). The prefix *ma-* is also used in the formation of agent and patient nominalisations in Gavar (6) (see also Viljoen 2017b: 46–49).

(6)a. ***mà****-ʔʷə́n* *dʒàk,* *tà-hà.*
NMLZ-build house arrive-VNT
'The builder has arrived.'

b. ***má****-báɮ-á-bàɮ* *à:gí* *xì*
NMLZ-exterminate-NMLZ.PAT-exterminate PL person
má *dəgnàk* *blə̀kʷ* *nfàɗ* *ábə*
REL.SBJ black thousand four with

DOI 10.15460/auue.2025.98.1.354

rà-nʃílíɗ *áxà* *dzánfáɗ.*
hand-seven over nine
'Four thousand and seventy-nine people were killed.' (lit. 'The ones killed were four thousand and seventy-nine.')

The similarity of the nominaliser and relative marker forms has also been noted for other Chadic languages such as Miya (Schuh 1998: 260–276), Buwal (Viljoen 2013: 298), Mofu-Gudur (Hollingsworth 2004: 43), Cuvok (Dadak 2021: 180) and Mina (Frajzyngier & Johnston 2005: 38–39, 429).

The tone on the Gavar subject relative marker *ma* varies according to its position in the clause, the tone on the preceding word and the word class of the word that follows it. This variation is somewhat unexpected as the tone on the relative marker in certain closely related languages appears to be stable, being always high in Buwal (Viljoen 2013: 298), and low in Mina (Frajzyngier & Johnston 2005: 429). However, examples of relative clauses given for Daba (Lienhard & Giger 1976: 15) show a variation between low and high on the relative marker *ma*. No explanation for this variation is given. From the examples, it seems that the tone on *ma* is high when preceded and followed by a high tone, and low when preceded and followed by a low tone.

Table 1 below gives a summary of the tonal variations observed on the relative marker *ma* for Gavar and where they occur. Examples follow below.

Table 1: Summary of tonal variations of the subject relative pronoun *ma*

Low	*mà* PREP
	[low tone word] *mà* VERB
	[no Head Noun] *mà* VERB (subject position)
High	*má* ADJ
	má NUM
	[high tone word] *má* VERB
	[no Head Noun] *má* VERB (non-subject position)

a) For non-verbal relative clauses, the word class of the following word determines the tone on *ma*. Preceding a preposition such as *átá* 'on', in example (7) below, the relative marker carries a low tone *mà*.

DOI 10.15460/auue.2025.98.1.354

(7) *límlím* [***mà*** *átá* *áŋɮìŋ*]RC
dew REL.SBJ on peanut
'the dew which is on the peanut plant'

Whereas preceding an adjective or numeral such as *dəgnàk* 'black' in example (8a) and *ntáɗ* 'one' in example (8b), the tone is high.

(8)a. *rək^{w}ət͡s* [***má*** *dəgnàk*]RC *tá-skəm-ájá.*
garment REL.SBJ black PFV-buy-PTCP
'The garment which is black has been bought.'

b. *xəjì-ká-dəm* *átá* *là* [***má*** *ntáɗ.*]RC
1INCL.SBJ-IPFV-sort on place REL.SBJ one
'We have sorted together.' (lit. 'We have sorted in the one place.')

b) For verbal relative clauses, the tone on the final syllable of the preceding word spreads onto the relative marker *ma*, so that if the preceding tone is low as on *àlà* 'someone' in example (9a) below, the tone on *mà* is low. If the preceding tone is high as on *mávə́* 'beer', the tone on *má* is high. This is similar to what was observed for Daba, as mentioned above.

(9)a. *àlà* [***mà*** *d͡za* *àlà*]RC *tà-ʃi* *tə.*
someone REL.SBJ kill someone PFV-flee 3SG.POSS
'The person who killed someone has fled.'

b. *mávə́* [*má* *gəj* *tə*]RC *wàtsà*
beer REL.SBJ spoil 3SG.POSS DEM.DIST
tá-sa-ájá.
PFV-drink-PCPT
'That beer which is spoilt has been drunk.'

c) For headless verbal relative clauses, its position in the clause determines the tone of the relative marker. If the headless RC is in subject position, as in examples (10a) below and (5a) above, *mà* carries a low tone. The tone is also low when the RC is in the predicate position following the focus copula *ànà* (10b).

DOI 10.15460/auue.2025.98.1.354

(10)a. [***mà*** *nda-xà*]RC *vàjà?*
REL.SBJ go-VNT who
'Who comes?' (lit. 'The one who comes is who?')

b. *xədʒì gbák ànà* [***mà*** *ndəv.*]RC
person two COP REL.SBJ fall
'It was two people who fell.'

If the headless relative clause is in a predicate position with no copula preceding, as in example (11) below, or another non-subject position such as direct object (5b) or indirect object (5c), the relative marker *má* carries a high tone.

(11) *là-táŋgʷál gèŋ* [***má*** *ká-ndrim-ì*
NMLZ.ACT-roll boulder REL.SBJ IPFV-hurt-3SG.DOBJ
á dəɓà.]RC
LOC back
'Rolling the boulder is what hurt him in the back.'

The non-subject relative marker has the form *ndə́* (free variant *ná)* and is used with all other grammatical relations apart from the subject. Unlike the tone on the subject relative marker, the tone on the non-subject pronoun is stable, always being high (12a–b). Furthermore, the tone on a following singular subject marker on the verb is also high, as in *sá-* 1SG.SBJ in example (12a) and *xʷá-* 2SG.SBJ in example (12b). In main indicative clauses the tone on corresponding singular subject markers is low, as in *sà-* 1SG.SBJ in example (13a) and *xʷà-* 2SG.SBJ in example (13b) (Viljoen 2019: 263).

(12)a. *rəkʷətsˀ* [***ndə́*** ***sá****-tsa-ə*]RC *ɓəzmək*
garment REL.NSBJ 1SG.SBJ-put-3SG.DOBJ dark
íkʷá.
like.this
'The garment that I wear is dark.'

b. *bìk* [***ndə́*** ***xʷá****-ká-ntir* *ábə* *ʃà*]RC
pen REL.NSBJ 2SG.SBJ-IPFV-write with 3SG
'the pen that you write with'

DOI 10.15460/auue.2025.98.1.354

(13)a. ***sà****-skəm-ɓà* *nxʷà.*
1SG.SBJ-buy-REFL.IOBJ goat
'I buy myself a goat.'

b. ***xʷà****-diʃ* *rà* *tə̀*
2SG.SBJ-touch arm 3SG.POSS
'You touch his arm.'

3.3 The strategy used to express NP_{rel}

The strategy used to express the NP_{rel} within the relative clause varies according to the grammatical role of the noun being relativised.

For subjects we see a combination of strategies being used. As was seen in the in previous section (3.2), a subject NP_{rel} is encoded with the subject relative pronoun *ma* within the relative clause. However, Gavar is a language with subject cross-referencing on the verb, even when an overt subject NP is present (14).

(14) *wálí* ***à:****-sa* *jàká* *dàgàlà* *átá*
women 3PL.SBJ-drink suffering a.lot on
là-mbə́w.
NMLZ.ACT-give.birth
'Women suffer a lot while giving birth.'

With regard to this subject agreement marking, a gap strategy is used, the expected prefix being omitted in subject relative clauses (15 a– b). The gap strategy is used regardless of the tense/aspect marking on the verb (see Section 3.5).

(15)a. *sà-gas* *wálí* [***má*** *ká-ŋtir* *skə̀n*
1SG.SBJ-look.for women REL.SBJ IPFV-write thing
à:gí.]RC
PL
'I am looking for women who write things.'

b. *ʃà* *ànà* [***mà*** *ká-á:-nda.*]RC
3SG COP REL.SBJ IPFV-FUT-go
'It is he who will be going.'

For non-subjects, a common strategy is that of pronoun retention (Payne 1997: 331) or the resumptive strategy (Tallerman 2015: 282),

DOI 10.15460/auue.2025.98.1.354

where a resumptive pronoun appears in the relativised position. However, in some cases a gap strategy is also possible (examples given below). For the direct object and indirect object, the resumptive pronoun takes the form of a verbal suffix, as in *-ə* 3SG.DOBJ in example (16a), and *-ànə* 3SG.IOBJ in example (16b).

(16)a. *tsákʷà ànà ɮən [ndə́ hwá-ká-ɗàl-**ə**.*]RC
here COP work REL.NSBJ 2SG.SBJ-IPFV-do-3SG.DOBJ
'Here is the work that you are doing (it).'

b. *wəʒí ànà [ndə́ sá-ʔʷəf-**ànə***[3]
children COP REL.NSBJ 1SG.SBJ-heat-3SG.IOBJ
wìrì.]RC
sauce
'It is the children for whom I heat the sauce.' (lit. 'It is for the children that I heat sauce (for them).')

While the verbs in the vast majority of direct object relative clauses in the data set carry a resumptive pronominal marker suffixed on the verb, there are a few examples, where it is omitted (17). It is unclear at this point whether this variation is free or conditioned in some way.

(17) *sà-tá-ŋga nkʷáb átá skən [ndə́*
1SG.SBJ-PFV-break brain on thing REL.NSBJ
sá-ká-ɗal gámák.]RC
1SG.SBJ-IPFV-do bad
'I took account of the bad things I do.'

For some obliques the resumptive pronoun takes the form of an independent pronoun functioning as the object of a preposition, as in example (18a), where the pronoun *ʃà* 3SG is the object of the preposition *ábə* 'with'. In other cases, the pronoun is incorporated within the preposition, as in *á xédé* 'on it' in (18b).

(18)a. *ántʃì [ndə́ sá-ká-nda **ábə** **ʃà**]*RC
shoe REL.NSBJ 1SG.SBJ-IPFV-go with 3SG
'The shoe that I am walking with (it).'

3 Note that frequently in natural speech the third person singular indirect object agreement marker is used when referring to a plural indirect object.

DOI 10.15460/auue.2025.98.1.354

b. *jí* *á-ji* *nkədàŋ* [*ndə́* *á-ká-tsa*
take 3SG.SBJ-take stone REL.NSBJ 3SG.SBJ-IPFV-put
xà *tə* ***á*** ***xédé.***[4]]RC
head 3SG.POSS LOC on.it
'He took the stone that he was laying his head on (it).'

It is possible, however, for the gap strategy to be used for temporal and locative oblique relative clauses, as in examples (19a) and (19b) below where the time and location are not encoded within the relative clause.

(19)a. *ntrà* [*ndə́* *ní-ká-dʒi* *xà* *tʃînì*]RC
month REL.SBJ 1EXCL-IPFV-take head 1EXCL.POSS
'the month that we are marrying each other' (lit. 'the month that we are taking our head(s)')

b. *váŋ*[5] *á-vaŋ* *tə̀* *átá* *là* [*ndə́*
arrive 3SG.SBJ-arrive 3SG.POSS on place REL.NSBJ
á-wan.]RC
3SG.SBJ-sleep
'He arrived at the place where he slept.'

A relativised possessor is encoded by a possessive adjective following the possessed noun within the relative clause, as in *tə̀* 3SG.POSS which follows *nx^{w}à* 'goat' in example (20) below.

(20) *àlà* [*ndə́* *nx^{w}à* ***tə̀*** *à-fəcak*
someone REL.NSBJ goat 3SG.POSS 3SG.SBJ-get.lost
tə̀]RC *wàtsə́,* *à-ká-xan.*
3SG.POSS DEM.PROX 3SG.SBJ-IPFV-cry
'This person whose (his) goat got lost, is crying.'

A gap strategy is used with regards to the object of comparison. In example (21) below, nothing follows the comparative preposition *áxà*.[6]

4 This form is also commonly found in main clauses.

5 A leftward repetition of the verb root results in an emphatic or highlighted verb form (Viljoen 2017b: 65–67).

6 See Section 4 for examples of non-relativised comparative clauses

DOI 10.15460/auue.2025.98.1.354

(21) *àlà* [*ndə́* *àlà* *wàtsə́* *pá* *má*
person REL.NSBJ person DEM.PROX at.level REL.SBJ
dʒíŋ ***áxà***]RC *wàtsá,* *à-ɮap* *tə́wáx* *íjkʷà.*
tall over DEM.DIST 3SG.SBJ-speak good like.this
'That person that this person is taller than, speaks well.'

3.4 Which grammatical relations can be relativised

In Gavar all the grammatical relations in the accessibility hierarchy (see Fig. 1), may be relativised (see examples (22) to (27) below).

a) Subject

(22) *ɮà* *tə̀* [*mà* *fəcàk* *tə̀* *gədə́k*]RC
ox 3SG.POSS REL.SBJ get.lost 3SG.POSS far
náná *tá-gəra-ájá.*
DEM.DIST PFV-find-PTCP
'That ox of his that got lost far away has been found.'

b) Direct object

(23) *kʷá* *skàn* *tə̀* *vémé* [*ndə́*
even thing 3SG.POSS what REL.NSBJ
á-ʔʷəs-ə̀]RC *à-ɗal* *tə́wáh.*
3SG.SBJ-cultivate-3SG.DOBJ 3SG.SBJ-do good
'Anything of his that he cultivates (it) does well.'

c) Indirect object

(24) *wəʒí* *ànà* [*ndə́* *sá-ʔʷəf-ànə̀* *wìrì.*]RC
children COP REL.NSBJ 1SG.SBJ-heat-3SG.IOBJ sauce
'It is for the children that I heat sauce (for them).'

d) Oblique

In the corpus, examples of relativised obliques encoding instrument (25a), location (25b) and time were found (25c).

(25)a. *ɓàrɮàm* [*ndə́* *xì-ká-ra* *jàm* *ábà*
iron REL.NSBJ 1INCL.DOBJ-IPFV-dig water with
ʃà]RC
3SG
'The iron (tool) that we dig the well with (it).'

DOI 10.15460/auue.2025.98.1.354

b. *dʒàk ànà* [*ndə́ wàlà à-ɮaɗ ŋgàlèlèm*
hut COP REL.NSBJ woman 3SG.SBJ-sweep gravel
á mə́n.]RC
inside
'It is the hut that the woman sweeps gravel from (inside).'

c. *xʷà-tá-ndʒin-ànì á vàh* [*ndə́*
2SG.SBJ-PFV-follow-1EXCL.DOBJ LOC *day* REL.NSBJ
ní-vah á màxʷə̀bə̀m kʷa.]RC
1EXCL.SBJ-pass.day LOC front 2SG.POSS
'You have watched over us during the day that we passed before you.'

e) Genitive

(26) *àlà* [*ndə́ rə̀kʷə̀ts tə̀ tá-jaɓ-ájá,*]RC
someone REL.NSBJ clothes 3SG.POSS PFV-clean-PCPT
à-ká-nda tə̀ á gʷàjgʷàjà.
3SG.SBJ-IPFV-go 3SG.POSS LOC festival
'The person whose clothes have been cleaned, is going to the festival.'

f) Object of comparison

(27) *ɮà* [*ndə́ nxʷà ànà má pá də̀j*
ox REL.NSBJ goat COP REL.SBJ at.level greater
áxà]RC *wàtsə́, à-ká-ʃi tə̀.*
over DEM.PROX 3SG.SBJ-IPFV-run 3SG.POSS
'This ox that the goat is bigger than, is fleeing'

3.5 Tense/aspect in relative clauses

Verbs in independent clauses in Gavar may be marked with perfective (*tá-*), imperfective (*ká-*), or future (*á:-*) prefixes (Viljoen 2019: 263). Verbs in Gavar are also frequently left unmarked for TAM. Unmarked verbs are used for generic situations or actions and can refer to past events, being used for theme-line, consecutive events in narrative texts (Viljoen 2019: 263–264).

In relative clauses, unmarked (9a), imperfective (15a) and future (28a) verb forms can be found. However, the perfective form is not

DOI 10.15460/auue.2025.98.1.354

permitted (28b). The perfective meaning is covered in relative clauses by the unmarked verb as in example (9a).

(28)a. *màwəl* [*mà* *áː-ʔʷən* *dʒàk,*]RC *tà-hà.*
man REL.SBJ FUT-build house arrive-VNT
'The man who will build the house has arrived.'

b. **màwəl* [*mà* *tá-ʔʷən* *dʒàk,*]RC *tà-hà.*
man REL.SBJ PFV-build house arrive-VNT
'The man who built the house has arrived.'

It is not unusual in Chadic languages for tense/aspect marking within relative clauses to differ in some way from that in independent clauses (Frajzyngier 1996: 454–455). For example, in Giziga (Shay 2021: 301), Mofu-Gudur (Hollingsworth 2004: 43), Cuvok (Dadak 2021: 183) and Sukun (Thomas 2014: 242), past tense marking is not attested in relative clauses but future marking is possible. This is reminiscent of the restriction on perfective marking in Gavar. The special treatment of the perfective in relative clauses can also be seen in West Chadic languages such as Hausa which has two perfective forms, one of which is only found in non-relative contexts, while the other occurs in both relative and non-relative contexts (Zima 1991: 16).

So-called 'relative' verb forms are common in both the West and East branches of Chadic (Zima 1991: 22). Certain Central Chadic languages also, for example Gidar (Frajzyngier 2008: 455–456) and Cuvok (Dadak 2021: 183–184), have special subordinate verb forms which are used in relative clauses. This, however, is not the case for Gavar, which apart from the restriction on the perfective, uses the same verb forms in relative clauses as are found in independent clauses. There is also no distinction in the form of the relative marker with regard to perfective/imperfective aspect in Gavar as may be found in certain other Central Chadic languages such as Giziga (Shay 2021: 300) and Daba (Lienhard & Giger 1976: 15f.).

3.6 Existential status of the head noun

The head noun of Gavar restrictive relative clauses can have either existential status. When its existence has not been previously established, there is no post-relative marker following the relative clause

DOI 10.15460/auue.2025.98.1.354

(29a–b). The head noun may either be non-specific (29a) or specific (29b).

(29)a. *àː-ká-daː-xà* *àlà* [*mà* *srək-ànə*
3PL.SBJ-IPFV-bring someone REL.SBJ teach-3SG.IOBJ
skàn *á* *wàʒí*]RC [*má* *ntágʷə́lèŋ.*]RC
thing LOC children REL.SBJ one
'They bring only one teacher.'

b. *ʔʷə̀làŋ* [*ndə́* *sá-ɓam-ə̀*]RC *rà-dzánfáɗ*
panther REL.NSBJ 1SG.SBJ-munch-3SG.DOBJ hand-nine
áxà *dzánfáɗ.*
over nine
'The panthers that I ate are ninety-nine.'

When the existential status of the head noun has already been established, a demonstrative follows the relative clause (22 & 30).

(30) *sà-tá-waj* *mà* *kʷà* [*ndə́*
1SG.SBJ-PFV-forget word 2SG.POSS REL.NSBJ
xʷá-ɣal-àkà]RC *wàtsà* *bàzà.*
2SG.SBJ-tell-1SG.IOBJ DEM.DIST COMPL
'I have forgotten what (lit. that word) you told me.'

Therefore, we see that Gavar encodes the existential status of the head noun by contrasting the use versus the absence of post-relative markers. This corresponds with Frajzyngier's (1996: 432) observation for some Central Chadic languages, that the only way of indicating the 'definiteness' of the head noun is by adding a post-relative marker. A similar system is found in the nearby language Mina (Frajzyngier & Johnston 2005: 430).

4 Adverbial clauses based on the relative clause structure

In Gavar, there are a number of different types of adverbial clauses which have developed from the relative clause structure. These differ from standard relative clauses in that they do not modify nouns, nor function as arguments in a clause. Rather they have an adverbial function. They all involve the non-subject relative marker *ndə́* (or *ná*). Some are marked directly with *ndə́* (Section 4.1), in other cases

DOI 10.15460/auue.2025.98.1.354

the relative marker is preceded by a preposition (Section 4.2). The marking of adverbial clauses with a relative marker has also been observed in nearby languages such as Daba (Lienhard & Giger 1976: 16), Mofu-Gudur (Hollingsworth 2004: 45) and Buwal (Viljoen 2013: 564–572).

4.1 Adverbial clauses directly marked with *ndə́*

There is some variation in semantic relations that may be expressed by adverbial clauses directly marked with *ndə́*. The most common, however, is the general temporal relation. One could hypothesise that such temporal adverbial clauses originated from an oblique relative clause involving a temporal noun modified by a relative clause as in example (31a). Over time, the temporal noun is omitted (32b) with the resulting headless relative clause being reanalysed as an adverbial clause.

(31)a. ***ɣàlà*** [*ndə́* *xì-gʷaɓ* *là* *bəzà,*]RC
moment REL.NSBJ 1INCL.SBJ-weed field COMPL
xì-nda *tàkʷà* *á* *bəl* *mpì.*
1INCL.SBJ-go 1INCL.POSS INF chop tree
'The moment we finish weeding the field, we go to chop trees.'

b. [*ndə́* *xì-gʷaɓ* *là* *bəzà,*]AC *xì-nda*
REL.NSBJ 1INCL.SBJ-weed field COMPL 1INCL.SBJ-go
tàkʷà *á* *bəl* *mpì.*
1INCL.POSS INF chop tree
'When we finish weeding the field, we go to chop trees.'

For general temporal adverbial clauses, the exact nature of the temporal relation can be determined by the tense/aspect markers in each clause. In example (32a) the completive marker *bəzà* in the first clause indicates that the action expressed by this clause occurred before the action expressed by the second clause. The imperfective prefix *ká-* attached to all the verbs in example (32b) indicates that the actions/situations occurred at the same time.

(32)a. [***ndə́*** *tsáx* *xì-tsax-xà* *là* *bəzà,*]AC
REL.NSBJ clear 1INCL.SBJ-clear-VNT field COMPL

DOI 10.15460/auue.2025.98.1.354

féŋ hì-fèŋ-xà là bàzà.
burn 1INCL.SBJ-burn-VNT field COMPL
'When we have finished clearing the field, we burn it.'

b. [***ndə́*** *á-ká-ɗa mà?^{w}às à dʒàk,*]RC
REL.NSBJ 3SG.SBJ-IPFV-be newborn LOC hut
à-ká-ɮap, à-ká-nda.
3SG.SBJ-IPFV-speak 3SG.SBJ-IPFV-walk
'When he was a newborn in the hut, he was speaking, he was walking.'

Many such temporal clauses also have a cause type meaning, as in examples (33a) and (33b). The verbal aspectual particles *ká* and *àkà* in the temporal clause in each example respectively, indicate that the action expressed by the first clause occurs before that of the second. The purposive marker *ká* indicates that an action is being done in advance with a purpose in mind. The immediate marker *àkà* indicates that an action has just taken place.

(33)a. [***ndə́*** *xì-ʃin-à-zə́ ká,*]AC *bə́w*
REL.NSBJ 1INCL.SBJ-grind-VNT-TRANS PURP change
á-bəw tə́ á nfá ànà ndrì.
3SG.SBJ-change 3SG.POSS LOC flour COP sorghum
'When we grind it, the sorghum turns into flour.'

b. [***ná*** *mɮà à-vah á tal*
REL.NSBJ blacksmith 3SG.SBJ-pass.day INF hammer
dəvər àkà dàgàlà,]AC *rà tə́ tá-ndrim ká.*
hoe IMM a.lot arm 3SG.POSS PFV-hurt PURP
'When the blacksmith has just spent the whole day hammering hoes, his arm hurt.'

When a clause marked only with *ndə́* occurs in the predicate position of a copula clause, preceded by the focus copula *ànà*, it may express reason (34a) or manner (34b).

(34)a. *ʃà ànà* [*ndə́ sá-ká-maj-ɓà*
3SG COP REL.SBJ 1INCL.SBJ-IPFV-love-REFL.IOBJ

DOI 10.15460/auue.2025.98.1.354

Kʷásmá Raʃel áxà Kʷásmá Salawme.[7]]AC
Kosma Rachel over Kosma Salome
'It is why I love Kosma Rachel more than Kosma Salome.'

b. *ʃà ànà* [*ndɔ́ násárá à:gí à:-nda-xà*
3SG COP REL.SBJ white.man PL 3PL.SBJ-go-VNT
á gra-ànì]AC *wàtsɔ́.*
INF find-1EXCL.DOBJ DEM.PROX
'It is how the white men came and found us.'

4.2 Adverbial clauses marked with a preposition plus *ndɔ́*

Many prepositions in Gavar are complex, being a combination of the basic locative preposition *á* plus another element, often a body part noun (Viljoen 2022: 459). Taking into account their nominal source, it is not surprising that constructions involving such prepositions followed by a relative clause should have developed over time. The nominal part of the complex preposition would have originally functioned as the head noun modified by the relative clause, before being reanalysed as a construction involving a preposition plus a headless relative clause. Such constructions express a number of adverbial meanings, which are listed below with examples.

Note, however, that it seems that the relative pronoun is only present when the adverbial clause is verbal. Non-verbal adverbial clauses are marked directly with the preposition as in examples (35a) and (35b).

(35)a. *á dámɔ́, kʷáhʷɔ́ á xɔ̀jì à:-gra là*
LOC bush fire COMP person 3PL.SBJ-see place
[***ándzá*** *tíw kuraŋ à:gí ákàs.*]AC
as ASS.PL electricity PL NEG.EXIST
'In the bush there is fire for people to be able to see, as there are no things like electricity.'

b. *à-gas-bà wàlà á dʒi,*
3SG.SBJ-look.for-REFL.IOBJ wife INF take
à:-ká-vəl-ànɔ̀ skʷɔ́ [***árá*** *màsàgàl*
3PL.SBJ-IPFV-give-3SG.DOBJ NEG because laziness

7 Tones have not been marked on the French proper nouns in this example.

DOI 10.15460/auue.2025.98.1.354

tə *dəj.*]AC
3SG.POSS too.much
'He looks for a wife to marry, they don't give him one because he is too lazy (lit. his laziness is too much).'

4.2.1 Clauses marked with á dìwʒá

Clauses marked with *á dìwʒá* 'behind, after', express a relative time meaning, with the event expressed by the main clause occurring after the event in expressed by the adverbial clause (36).

(36) [***á dìwʒá*** ***ndə́*** *xì-gʷaɓ* *là* *bəzà,*]AC
after REL.NSBJ 1INCL.SBJ-weed field COMPL
xì-nda *tàkʷà* *á* *bəl* *mpì.*
1INCL.SBJ-go 1INCL.POSS INF chop tree
'After we have finished weeding the field, we go to chop trees.'

It is possible for words signifying time such as *vàx* 'day' to be inserted between *á dìwʒá* 'after' and the non-subject relative marker *ndə́* (37).

(37) [***á dìwʒá*** ***vàx*** ***ndə́*** *xì-gʷaɓ* *là*
after day REL.NSBJ 1INCL.SBJ-weed field
bəzà,]AC *xì-nda* *tàkʷà* *á* *bəl* *mpì.*
COMPL 1INCL.SBJ-go 1INCL.POSS INF chop tree
'After the day that we have finished weeding the field, we go to chop trees.'

4.2.2 Clauses marked with ándzá

Clauses marked with the preposition *ándzá* 'like, as' express manner (38).

(38) *à-ɗal* *á* *wàtà* *màwəl* *tə* [***ándzá***
3SG.SBJ-do LOC home husband 3SG.POSS as
ndə́ *á-ká-ɗal* *á* *wàtà* *bàbà* *tə.*]AC
REL.NSBJ 3SG.SBJ-IPFV-do LOC home father 3SG.POSS
'She does at her husband's home just as she was doing at her father's home.'

DOI 10.15460/auue.2025.98.1.354

4.2.3 Clauses marked with árá

Clauses marked with the preposition *árá* 'to, at (a person), because' express reason (39).

(39) *màɬíjí nàkà à-ká-tab wəlàf*
neighbour 1SG.POSS 3SG.SBJ-IPFV-touch blindness
[***árá ná*** *ndʒà tə*
because REL.NSBJ eye 3SG.POSS
à-ká-ndrim-ì.]AC
3SG.SBJ-IPFV-hurt-3SG.DOBJ
'My neighbour has difficulty seeing because his eyes hurt him.'

4.2.4 Clauses marked with átá

Clauses marked with the preposition *átá* 'on, about' can refer to a discourse topic (40).

(40) *ʃà ànà mà ndə́ sá-ká-ɣal-axʷə*
3SG COP word REL.NSBJ 1SG.SBJ-IPFV-tell-2SG.IOBJ
[***átá ndə́*** *sá-ká-maj Kʷásmá.*]AC
about REL.NSBJ 1SG.SBJ-IPFV-love Kosma
'These are the reasons that I was telling you about why I love Kosma.'

There was also the following example in the corpus (41) which seems to express the evidence from which one may deduce a fact.

(41) *là wár, xì-sən* [*átá ndə́*
havest.season 1INCL.SBJ-know on REL.NSBJ
ndrí á-ká-kpaj.]AC
sorghum 3SG.SBJ-IPFV-flower
'The harvest season, we know it from the fact that the sorghum is flowering.'

4.2.5 Clauses marked with áská

The preposition *áská* 'under, for, so that' marks purpose clauses in Gavar. In the majority of cases a relative marker is not present. However, there are two examples of this preposition in the corpus being followed by the non-subject relative pronoun (41).

DOI 10.15460/auue.2025.98.1.354

(42) *sà-tá-kəɗaŋ ra* [***ásḱá*** ***ná***
1SG.SBJ-PFV-finish SIM so.that REL.NSBJ
sá-ká-xaʃaʃaŋ bəzà.]AC
1SG.SBJ-IPFV-forget COMPL
'I finished on the way lest I forget'

5 Comparative clauses

Certain comparative constructions in Gavar are based on the subject relative clause with an adjectival predicate and marked with the subject relative pronoun carrying a high tone *má.* This reflects the parameters under Table 1 as it is seen there that the relative marker carries a high tone when followed by an adjective. The relative marker is usually preceded by the preposition *pá* 'at level' while the object of comparison is preceded by the preposition *áxà* 'over' (43).

(43)a. *mpì* ***pá*** [***má*** *ʤíŋ áxà ndrì.*]RC
tree at.level REL.SBJ tall over sorghum
'The tree is taller than the sorghum.'

b. *xàjàk nàkà* ***pá*** [***má*** *dĭmdĭm áxà*
country 1SG.POSS at.level REL.SBJ cool over
ntrà nkrə̀m.]RC
month dry
'My country is cooler than the dry season.'

One example was found where the preposition *pá* was not present. In this case the relative clause was directly preceded by the focus copula *ànà* (44).

(44) *K^{w}àjàŋ ànà* [***má*** *dèf áxà Làwán.*]RC
Koyang COP REL.SBJ short over chief
'Koyang is shorter than the chief.'

Superlatives have a similar structure but in this case the object of comparison is not explicitly mentioned as in examples (45a) and (45b). Note that the preposition *áxà* is optional. It is present in example (45b), while in (45a) it is missing.

DOI 10.15460/auue.2025.98.1.354

(45)a. *míndʒá* ***pá*** [***má*** *łéx.*]RC
DET.INDF.SG at.level REL.SBJ very.thin
'The other one is the thinnest.'

b. *mbə̀* *gàmtàk* ***pá*** [***má*** *féŋ* *áxà.*]RC
child chicken at.level REL.SBJ small over
'The chick is the smallest.'

6 Conclusion

Embedded restrictive relative clauses are used extensively in Gavar and exhibit many characteristics common to relative clauses in other languages. Like most other VO languages (Payne 1997: 326; Tallerman 2015: 278), and many other Chadic languages, the Gavar relative clause follows its head noun which is external to it. The relative clause is marked by a relative marker at the beginning of the clause. Gavar, in common to several nearby Central Chadic languages, has two distinct relative markers; one for the subject role and one that covers all non-subject roles. It is interesting to note that the tone on the subject relative marker *ma* is variable, while the tone on the non-subject relative marker *ndə́ (ná)* is stable. Such tonal variation is not frequently reported for other Chadic languages in the area. The distinction of subject and non-subject carries over into how the NP_{rel} is expressed within a relative clause. The gap strategy is used for subjects with regard to subject agreement marking on the verb, while resumptive strategy is most common for non-subjects. This is a common pattern cross-linguistically. Unlike many of the world's languages, Gavar is able to relativise all grammatical functions found in the hierarchy of grammatical relations: subjects, direct objects, indirect objects, obliques, genitives and objects of comparison.

Gavar does not make use of the 'relative tenses' that are commonly found in the West and East branches of Chadic. There is, however, a restriction on perfective marking in relative clauses which corresponds with similar restrictions observed in nearby languages. The existential status of the head noun is coded by the absence versus the presence of a post-relative demonstrative, a strategy frequently seen in Chadic languages (Frajzyngier 1996: 431–433).

DOI 10.15460/auue.2025.98.1.354

As in many languages, other complex constructions in Gavar such as adverbial clauses and comparatives have developed over time from relative clause constructions. Adverbial clauses involve the use of the non-subject relative marker. This is not surprising as they most likely have developed from oblique relative clauses. The relative marker may mark the adverbial clause alone in which case the most common meaning is a general temporal one whose interpretation depends of the tense/aspect marking of the verbs. The relative pronoun may also be preceded by a number of different prepositions yielding more specific semantic relations such as: temporality, manner, reason, discourse topic and even purpose clauses. Comparatives (and superlatives), on the other hand, are based on subject relative clauses with an adjectival predicate.

Further research needs to be done into the possibility of non-restrictive as well as adjoined relative clauses. Also reduced relative clauses (Andrews 2007: 232–233) involving such participle forms as *dzawájá* 'tied up' in example (46) below, warrant investigation.

(46) *nxʷà mà dzaw-ájá, tá-mtał tə̀.*
goat REL.SBJ attach-PTCP PFV-break 3SPOSS
'The goat that was tied up, broke away.'

Abbreviations

1	first person	LOC	locative
2	second person	NMLZ	nominaliser
3	third person	NEG	negation
AC	adverbial clause	NP	noun phrase
ACT	action	NPREL	relativised noun phrase
ASS	associative	NSBJ	non-subject
COMP	complementiser	PAT	patient
COMPL	completive	PCPT	participle
COP	copula	PFV	perfective
DEM	demonstrative	PL	plural
DET	determiner	POSS	possessive
DIST	distal	PROX	proximal
DOBJ	direct object	PURP	purposive
EXCL	exclusive	RC	relative clause

DOI 10.15460/auue.2025.98.1.354

EXIST	existential	REFL	reflexive
HN	head noun	REL	relative pronoun
ICP	intransitive copy pronoun	SBJ	subject
IMM	immediate	SG	singular
INCL	inclusive	SV(O)	subject, verb, (object)
INDF	indefinite	TAM	tense, aspect, mode
INF	infinitive	TRANS	transitivity marker
IOBJ	indirect object	VNT	ventive
IPFV	imperfective aspect	VO	verb-object

References

Andrews, Avery D. 2007. Relative Clauses. In Timothy Shopen (ed.), *Language typology and syntactic description: Complex constructions*, vol. 2. 2nd edn. Cambridge: Cambridge University Press. 206–236. https://doi.org/10.1017/cbo9780511619434.004

Dadak, Ndokobai. 2021. *Grammaire cuvok: langue tchadique centrale du Cameroun*. Amsterdam: LOT.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2024. Gavar. *Ethnologue: Languages of the World.* Dallas: SIL International. http://www.ethnologue.com.

Frajzyngier, Zygmunt. 1996. *Grammaticalization of the complex sentence: A case study in Chadic.* (Studies in Language Companion Series 32). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1353/lan.1998.0251

Frajzyngier, Zygmunt. 2008. *A grammar of Gidar.* Frankfurt: Peter Lang.

Frajzyngier, Zygmunt & Eric Johnston. 2005. *A grammar of Mina.* Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110893908

Gravina, Richard. 2014. *The phonology of Proto-Central Chadic.* Utrecht: LOT.

Greenberg, Joseph. 1966. *The languages of Africa.* Bloomington: Indiana University.

Hollingsworth, Kenneth. 2004. *A survey of clause types in Mofu-Gudur.* Yaounde: SIL Cameroon. https://www.silcam.org/resources/archives/47713

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99. https://www.jstor.org/stable/4177973.

DOI 10.15460/auue.2025.98.1.354

Lienhard, Ruth & Marti Giger. 1976. *Daba grammar phrase level sorting work paper*. Yaounde: SIL Cameroon. https://www.silcam.org/resources/archives/47477

Noukeu, Serge. 2004. *Etude phonologique du Gavar*. Yaounde: University of Yaoundé I Maitrise.

Payne, Thomas. 1997. *Describing morphosyntax. A guide for field linguists*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511805066.012

Schuh, Russel G. 1998. *A grammar of Miya*. Berkeley & Los Angelos: University of California Publishing.

Shay, Erin. 2021. *A grammar of Giziga*. Leiden: BRILL. https://doi.org/10.1163/9789004445970

Tallerman, Maggie. 2015. *Understanding syntax*. 4th edn. London: Routledge. https://doi.org/10.4324/9781315758084

Tchikoua, Paul. 2006. *Morphologie nominale du Gavar*. Yaounde: Yaoundé I Maitrise en Linguistique Générale.

Thomas, Michael F. 2014. *A grammar of Sakun (Sukur)*. Boulder: University of Colorado.

Thompson, Sandra, Robert Longacre & Shin Ja Hwang. 2007. Adverbial Clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2. 2nd edn. Cambridge: Cambridge University Press. 237–300. https://doi.org/10.1017/cbo9780511619434.005

Trask, Robert Lawrence. 1993. *A dictionary of grammatical terms in linguistics*. London: Routledge. https://doi.org/10.4324/9780203393369

Viljoen, Melanie. 2013. *A grammatical description of the Buwal language*. Bundoora: La Trobe University. http://arrow.latrobe.edu.au:8080/vital/access/manager/Repository/latrobe:37216.

Viljoen, Melanie. 2017a. A comparison of the structural and functional properties of the optional copula in Buwal and Gavar. In Shigeki Kaji (ed.), *Proceedings of the 8th World Congress of African Linguistics, Kyoto 2015*. Kyoto: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies. 149–168.

Viljoen, Melanie. 2017b. *Gavar verb phrase*. Yaounde: SIL Cameroon. https://cameroon.sil.org/resources/archives/78579.

Viljoen, Melanie. 2019. Gavar verb morphology. In Henry Tourneux & Yvonne Treis (eds.), *Topics in Chadic linguistics X: Papers from the 9th Biennial International Colloquium on the Chadic Languages, Villejuif, September 7–8, 2017*. (Chadic Linguistics / Linguistique Tchadique / Tschadistik 11). 257–274.

DOI 10.15460/auue.2025.98.1.354

Viljoen, Melanie. 2022. Varying functions of the comitative preposition ábə in Gavar. *Language Typology and Universals* 75(3). 449–495. https://doi.org/10.1515/stuf-2022-1059.

Zima, Petr. 1991. Le problème des formes verbales dites 'relatives' en tchadique. In Herrmann Jungraithmayr & Henry Tourneux (eds.), *Études tchadiques: la phrase complexe*. (Études Tchadiques). Paris: Groupe d'études tchadiques. 15–26.

DOI 10.15460/auue.2025.98.1.354

The (low) phylogenetic relatedness of Conative Animal Calls: The case of Mokpe and Oroko (Bantu Zone A)

Pius W. Akumbu[a], Alexander Andrason[b] & Levi E. Mokake[c]

CNRS-LLACAN, Villejuif[a]
Living Tongues Institute for Endangered Languages, Salem & University of Cape Town[b]
University of Bamenda[c]
pius.akumbu@cnrs.fr
aleksand@hi.is
levimokake@gmail.com

Abstract
The present article is dedicated to the documentation of conative animal calls (CACs) in Mokpe and Oroko (Bantu Zone A) and the analysis of the cognancy / phylogenetic relatedness of these constructions. The examination of original data demonstrates that the CAC categories in both languages comply with synchronic and diachronic tendencies characterizing CACs in the world's languages. Primary CACs closely match the prototype of CACs and can be regarded as its canonical instantiations; secondary CACs may violate the prototypical features, which is also regular from a typological perspective. Furthermore, the comparison of Mokpe and Oroko CACs reveals the low cognancy of CACs in the two languages. The greater similarity exhibited by primary tokens than is the case of secondary ones does not stem from shared ancestry, but is more likely due to parallel developments and the exploitation of fully motivated, typologically common strategies.

Keywords: Human-to-animal communication, conative animal calls, interactives, phylogenetic relatedness, Bantu

1 Introduction

The present article is dedicated to the study of conative animal calls (CACs) in Mokpe ([bri]/[mokp1239]/[A22])[1] and Oroko ([bdu]/

1 The information provided in square brackets refers to ISO 639-3 codes, glottocodes, and Guthrie's (1967/1971) classification system in Hammarström's (2019) revision.

DOI 10.15460/auue.2025.98.1.347

[orok1266]/[A101]). We examine the similarity of CACs found in the above-mentioned varieties and discuss the potential phylogenetic (genealogical) relationship that underlies these types of constructions or their cognancy.

Mokpe and Oroko are Bantu languages of the so-called Sawabantu group, which comprises several zone A varieties: mainly A.20 and A.30, as well as, to some extent, A.10 (Nurse & Philippson 2003: 170, Ebobissé 2015). In the referential classification of the Bantu languages, Mokpe is part of the Duala group, i.e., A.20 (Hammarström 2019: 19) and the number of its speakers ranges from 25,000 to 35,000 people (Mutia 2005: 217–218, Atindogbe 2013: 5–6, Eberhard, Simons & Fennig 2023; see also Ebobissé 1989 and Neh 1989). Oroko forms part of the Lundu-Balong group, i.e. A.10 (Hammarström 2019: 19) and, according to the data from the beginning of this century, has between 110,000 and 140,000 speakers (D. Friesen 2002: 5; see also Atta 1993 and Eberhard, Simons & Fennig 2023). Mokpe and Oroko are closely related phylogenetically and, within the above-mentioned Sawabantu group, jointly belong to the greater Manenguba, Bafaw-Balong-Manenguba, and ultimately Bafawic-Bakweric branches. Within this last sub-branch, Mokpe is a Bakweric variety while Oroko is a Bafawic one (Eberhard, Simons & Fennig 2023, Hammarström et al. 2023; see also Ebobissé 1989, Neh 1989, Atta 1993, D. Friesen 2002).

Mokpe and Oroko are also geographically adjacent: the northern part of the Mokpe area connects to the southern part of the Oroko area (D. Friesen 2002, Atindogbe 2013). To be exact, Mokpe is spoken in the Wouri estuary and around the slopes of Mount Cameroon and Mount Fako, in cities such as Buea, Tiko, Limbe, and Muyuka in Western Cameroon (Atindogbé 2013, Tanda & Ayu'nwi 2005, Mokake 2016). Oroko is also spoken at the feet of Mount Cameroon but towards the north-west. Further Oroko settlements are found in the Meme and Ndian Divisions (Atta 1993, D. Friesen 2002, Mabian 2010, Che 1993, Atindogbé 2012). Both Mokpe and Oroko are in the Anglophone zone of Cameroon.

Conative animal calls (henceforth referred to as CACs), which constitute the topic of our article, are operationally defined as (a) directive lexemes or lexicalized constructions that (b) are addressed to animals and (c) can be (and, in fact, predominantly are) employed holophrastically as self-standing utterances or sentence equivalents

DOI 10.15460/auue.2025.98.1.347

(Andrason & Karani 2021, Andrason 2022, Andrason & Phiri 2023, Heine 2023, see also Bynon 1976, Aikhenvald 2010, Amha 2013). One of the issues recently observed in CAC scholarship concerns phylogenetics: CACs arguably exhibit a low level of relatedness or cognancy than what is attested in other lexical classes within a single linguistic family or branch (Daković 2006, Andrason & Akumbu 2024, Duah, Andrason & Antwi 2023, Andrason, Onsho Mulugeta & Shimelis Mazengia 2024, Andrason & Gafatsi 2025). In the present paper, we aim to contribute to the study of family-related properties of CACs and answer the following research questions: Do Mokpe and Oroko data confirm the phylogenetic hypothesis put forward on the basis of (very few) other languages and demonstrate the low cognancy of CACs? Should this be the case, we will respond to two further sub-questions: Are all types of CACs equally resilient to be shared in Mokpe and Oroko? What are the reasons of similarities, should these be attested? Additionally, given that Mokpe and Oroko are (heavily) under-researched, our article aims to contribute to the documentation of these two varieties.[2]

Following the most recent studies on CACs, including our own works dedicated to this lexical class (Andrason & Karani 2021, Andrason 2022, 2023, Andrason & Phiri 2023, Andrason & Akumbu 2024), the analysis proposed in this article will be guided by a prototype approach to linguistic categorization (Evans & Green 2006, Croft & Cruse 2004, Janda 2015) and the description of linguistic data will be developed with the terminological apparatus typical of less formal theories of language (which are commonly used in linguistic typology and from which our take on prototype theory emerges; see Croft 2003, Evans & Green 2006, Croft & Cruse 2004, Velupillai 2012, Janda 2015).

The article is organized as follows: in Section 2, we explain the theoretical background of our study. In Section 3, we introduce original evidence from Mokpe and Oroko and describe the function and

2 Nevertheless, important advancement in the documentation and analysis of these two languages have also been made. For Mokpe, these include studies by Hombert (1973), Gensler (1980, 1981a, 1981b), Hawkinson (1986), Bate (1987), Ebobissé (1989), Neh (1989), Kagaya (1993a, 1993b, 1993c), Connell (1997), Ngoisah Lyonga (2002), Marlo & Odden (2007), and Atindogbe (2013). For Oroko, less researched than Mokpe, the most important works include Atta (1993), Mbongue (2000), Friesen & Friesen (2001), D. Friesen (2002), and L. Friesen (2002).

DOI 10.15460/auue.2025.98.1.347

form of CACs as well as their relatedness/similarity. In Section 4, we evaluate our findings, answer the research question(s), and explain the contribution of our research to the scholarship of CACs. In section 5, we provide a conclusion to the study.

2 Theoretical background: CACs in related languages

CACs have traditionally been one of the most poorly researched categories in linguistics (Poyatos 2002: 178, Andrason & Karani 2021: 4–5). Indeed, it is only recently that their documentation and analysis have captivated more scholarly attention – a phenomenon that has its roots in the recognition of the relevance of interactives (e.g. interjections, ideophones, and directives) in language (cf. Heine 2023) and multimodality in human communication. This interest in CACs has led to a continuous expansion of typological evidence, proliferation of works examining these constructions in individual language varieties and language clusters, and much-improved theorizations regarding the form and function of a CAC category in human language, more generally.

To be exact, almost half a century ago, Bynon (1976) and Siatkowska (1976) offered the first comprehensive and theoretically advanced analyses of CACs.[3] The former concerned a Berber variety of Ayt Hadiddu, while the latter was dedicated to West Slavonic languages: Polish, Czech, and Slovak. 30 years later, research on CACs re-emerged thanks to Wierzbicka (2003), Daković (2006), Abdulla & Talib (2009), Amha (2013) and Denisova & Sergeev (2015), who reported more data from Polish, South and East Slavonic (i.e. Serbian, Croatian, and Russian), Arabic, Ta-Ne-Omotic (especially Zargulla) and Chuvash, respectively. Now, in the third decade of the 21st century, a new wave of studies, carried out primarily on the African continent, has introduced CACs to the forefront of language science. These recent research activities, which we will cite extensively in this paper, have systematically examined CACs in a Maasai language – Arusa (Andrason & Karani 2021); a Kwa language – Akan (Duah, Andrason & Antwi 2023); a Bantu language – Xhosa

3 Of course, these were not the first publications that mentioned CACs. See, for instance, Schulthess (1912), Jarring (1941), Eren (1952), and Germanovič (1954). For a review of the development of the research and scholarship dedicated to CACs consult Andrason & Karani (2021) and Andrason (2023).

DOI 10.15460/auue.2025.98.1.347

(Andrason 2022), Grassfields Bantu languages – Babanki, and Bum (Andrason & Akumbu 2024); and Oromo, Dogon and Khoe varieties, i.e. Macha Oromo (Andrason, Onsho Mulugeta & Shimelis Mazengia 2024), Togo-Teŋu Kan of Dourou (Andrason & Sagara 2024), and Tjwao (Andrason & Phiri 2023), respectively. Recently, Bernd Heine (2023) included CACs – referred to as 'animal directives' – as one of the categories of interactive grammar and, largely drawing on some of the above-mentioned studies, offered their detailed and compelling analysis.[4]

One of the most significant results of this recent proliferation of research on CACs is the design of the prototype of CACs that was put forward by Andrason & Karani in 2021 and subsequently revised by a number of other scholars (see, *inter alia*, Andrason & Phiri 2023, Andrason & Akumbu 2024, Andrason, Onsho Mulugeta & Shimelis Mazengia 2024, and Andrason & Sagara 2024). A prototypical CAC is defined by a series of semantic, pragmatic, phonetic, morphological, syntactic, and ecolinguistic properties. A detailed discussion of all such prototypical features may be found in the above-mentioned works. Below, we summarize those that are relevant for the present study: ecolinguistic, semantic, phonetic, and morphological properties. Ecolinguistically, both the form and function of CACs are heavily conditioned by their natural (e.g. topography, climate, fauna, and flora) and cultural (e.g. economy and social norms) context. Semantically, CACs typically express actions related to motion (summonses make animals to come, dispersals chase them away, and directionals modify their movement in any other way), have domestic species (rather than wild specifies) as their addresses, and are monosemous (instead of being polysemous). Phonetically, a prototypical CAC is a monosyllable or a chain of replicated monosyllables, exhibits consonantal (rather than vocalic) character, and hosts extra-systematic sounds and/or sound combinations, including suprasegmental, phonational, and "modulational" ones. Morphologically, CACs tend to be lexically opaque and monomorphemic, thus failing to make use of inflectional, derivational, and compounding strategies (alternatively, CACs consist of segments that, although replicated, are not genuine independent meaning-bearing units; in such cases, replication constitutes an expressive phonetic mechanism instead of a derivational

4 CACs have recently attracted attention of scholars working on other languages, e.g. Jääskeläinen (2021) on Finnish and Treis (2023) on Kambaata.

DOI 10.15460/auue.2025.98.1.347

morphological one). It is important to note that the formal properties enumerated above, both phonetic and morphological, tend to be exhibited by primary CACs (i.e. those that have entertained the status of CACs since their formation or have acquired it via grammaticalization and/or lexicalization to the extent that any non-CAC origin is no longer recoverable) but violated by secondary CACs (i.e. those that are derived from and still visibly coincide with other lexical classes).

While most of the studies mentioned above are language-specific and/or aim to contribute to the synchronic typology of CACs by designing, testing, and refining the prototype of a CAC, they have also made available some comparative and phylogenetic evidence. This evidence has, in turn, allowed linguists to formulate certain diachronic family-related hypotheses. Below, we chronologically enumerate the (certainly, very few) main proposals concerning the diachrony and phylogenetics of CACs put forward over the last half century:

a) In Slavonic languages (i.e. Croatian, Polish, Russian, and Serbian), cognate CACs are uncommon. Inversely, CACs that differ in form are more frequent (Daković 2006: 136, 142–144). Among all sub-types of CACs, those that are secondary or borrowed, as well as those that exhibit forms identical or similar to primary interjections are shared more commonly than CACs that are primary (Daković 2006: 143; see also Siatkowska 1976: 169–171).
b) The above observation was subsequently tested in a study dedicated to three closely related varieties of Akan: Asante, Bono, and Fante (Duah, Andrason & Antwi 2023). Akan data confirm Daković's (2006) proposal and demonstrate that (i) the cognancy and phylogenetic relation of CACs found in Akan, Bono, and Fante is much less visible than what characterizes the general word stock and core grammar and, thus, (ii) CAC constructions that could be inherited from a common Akan ancestor are much fewer than one would expect. This conclusion draws on two main facts. First, shared CACs are very few. This especially holds true of primary CACs, whereas the similarity of secondary CACs and the CACs that formally overlap with primary interjections is slightly more visible. Second, in most instances of formal resemblance, the similarity attested is not necessarily phylogenetic or inherited. Equally likely, it has

DOI 10.15460/auue.2025.98.1.347

its roots in separately exploiting the same strategies (onomatopoeic and non-onomatopoeic) to coin analogous constructions in the related varieties. Overall, although CAC categories in Asante, Bono, and Fante are canonical by largely complying with the prototype of CACs, they achieve this canonicity independently rather than inheriting it from a common ancestor.

c) The observations made with regard to the Akan language cluster have afterwards been corroborated by Ewe – another Kwa language (Andrason & Gafatsi 2025). In Ewe, like in Asante, Bono, and Fante, the category of CACs is canonical and matches the behavior associated with a prototypical CAC. However, CACs that are identical or similar in Ewe and Akan are few and, in their majority, do not descend from a shared ancestor. Rather, this synchronic similarity is due to areal phenomena (borrowing) or, even more likely, reflect common cross-linguistic strategies whereby the formal and functional similarity of some CAC lexemes arose in these two languages independently.

d) This "minimal" extent of cognancy and phylogenic affinity between related languages has subsequently been observed in Babanki and Bum, two Central Ring Grassfields Bantu varieties of Cameroon (Andrason & Akumbu 2024). With a few exceptions, Babanki and Bum lack CAC cognates in contrast with the general word stock where circa 85% of lexemes are cognate and descend from proto-forms. As in Asante, Bono, Fante, and Ewe, in cases where Babanki and Bum CACs exhibit analogous forms and functions, their similarity need not be phylogenetic (inherited) but may rather reflect crosslinguistic pressures. Indeed, in all such instances, comparable CAC constructions exist in many other languages unrelated to these two Grassfields varieties.

e) Lastly, some anecdotal data from Oromoid (Cushitic) languages, Macha Oromo and Faaʃe Konso, corroborate the diachronic/phylogenetic findings presented in points (a)–(d) above, as the CACs attested in these two varieties tend to be "distinct and etymologically unrelated" (Andrason, Onsho Mulugeta & Shimelis Mazengia 2024: 205).

Overall, the above-mentioned studies dedicated to CACs in Slavonic, Kwa, Bantoid, and Oromoid languages, suggest the following: (a) in related languages, cognate CACs are few; (b) more specifically, cognate CACs are fewer than what characterizes other lexical classes and

DOI 10.15460/auue.2025.98.1.347

grammatical categories; (c) when attested, the examples of similarity are better explained in terms of areal phenomena or independent/parallel developments; and, therefore, (d) CACs seem more resistant to be shared across (the history of) a family or branch of languages than is the case of many other types of words.

3 Evidence: CACs in Mokpe and Oroko

The present study of CACs in Mokpe and Oroko draws on data that were collected through introspection (one of the authors being a Mokpe native speaker) and semi-structured open-ended interviews. To recruit participants, we relied on the non-probabilistic discriminative snowball (network/chain) sampling technique. This data collection strategy consists of recruiting new participants thanks to the referral given to the researcher by current participants already enrolled and/or interviewed (Saumure & Given 2008, Eide 2008). This also allowed us to pre-screen participants following their age, knowledge of the language, and the duration of contact with the language. Altogether, we obtained data from nine Mokpe and three Oroko speakers.[5] The data collection process was conducted following an interview guide that had been developed for and used in our previous work dedicated to CACs in Babanki and Bum (Andrason & Akumbu 2024). The fieldwork took place between June and July 2023. The interviews were conducted in person or, given the ongoing Anglophone conflict in Cameroon, remotely via WhatsApp and direct telephone calls. The discussions were recorded with smartphones as .gpp or .acc audio files.

The data collection process described above permitted us to identify 43 CACs in Mokpe and 48 in Oroko. The selected CACs may be viewed as (relatively) stabilized in the respective speech communities.[6] Inversely, we did not include in our paper those CACs that seemed to be idiolectal or *hapax legomena*. Since the size of CAC categories in the languages studied thus far tends to oscillate between some 40 and 60 constructions (cf. Bynon 1976, Amha 2013,

5 The different number of consultants sampled per community resulted from the greater willingness to participate in the study expressed by Mokpe speakers as opposed to Oroko speakers.

6 However, as we explain below, the lexicalization of several secondary CACs is low.

DOI 10.15460/auue.2025.98.1.347

Andrason & Karani 2021, Andrason & Akumbu 2024, Andrason & Sagara 2024), we are confident that our data are *grosso modo* representative (although not exhaustive) of the two coastal Bantu languages studied in this article and allow for their meaningful comparison. Of course, one must bear in mind that we obtained data by interviewing only a handful of speakers. Although all CACs analyzed in this article are panlectal rather than idiolectal, it is possible that other Mokpe and/or Oroko speakers, e.g. people who hunt regularly, have slightly distinct repertoires of CACs. This however has no critical bearing on our study and its findings.

In Table 1, we list all the collected/selected CACs using the International Phonetic Alphabet (rather than the Mokpe and Oroko orthographies, cf. D. Friesen 2002 and Atindogbe 2013), with the exception of kisses, snaps, whistles, and CACs produced with objects, the highly extra-systematic realization of which will be described in detail in section 3.1. Additionally, we indicate the function of each token: we specify the action that a CAC is aimed to trigger and the animal kind to which it is (typically) directed.

Table 1. Conative animal calls in Mokpe and Oroko

Mokpe		**Oroko**	
Form	**Function**	**Form**	**Function**
ámbélɛ̂	request dogs to exercise patience when fed	*àráwâ*	chase away dogs
àŋdʒíjá áŋgâ	chase away pigs, cats, dogs	*árá wàgíjâ*	chase away cats, chickens
bâk	request dogs to go back	*bɛ̌ɛ́ɛ̀*	summon goats
βɛ́βɛ́lɛ̂	incite dog to chase game during hunting	*èràkà*	chase away pigs
dʒă nù	summon dogs	*érák(á) ò wókà*	request dog to go out of the house
dʒànâ	request dog to bring game during hunting	*éráká ùndáwùnɛ̂*	chase cats, chickens out of the house

DOI 10.15460/auue.2025.98.1.347

Mokpe		**Oroko**	
Form	**Function**	**Form**	**Function**
gō áùt	request dog to go out of the house	*éráká wètíndénɛ̂*	chase away goats
hôhôhô	chase chickens away	*fííí*	summon cats
hr̃ɔ̃ɔ̃	summon pigs	*gbâ*	incite dogs to chase game during hunting
ínsáì	request dogs, pigs to enter cage or sty	*gbɛ́ ò ndáwò*	request dogs to go into the house
kám yéè	request dogs to move towards speaker	*hr̃ɔ̃*	summon pigs
kɔ̀kɔ̀kɔ̀	summon chickens	*ìràkà nàŋgà*	instruct dogs to lie down
kùrúkùrú	summon chickens	*jáká*	summon dogs
líf dá plēs	chase away cats, dogs	*jáká wàŋgjâ*	summon goats
lèmbɛ̂	incite dogs to chase game during hunting	*kǎʃà*	incite dog to chase game during hunting
lìjâ	request dogs to sit down	*kátʃàm*	incite dog to chase game
mbólì	summon goats	*kómá*	incite dogs to chase game during hunting
mɛ̌ɛ́ɛ̀?	summon goats	*kɔ̀kɔ̀kɔ̀*	summon chickens
míjǎw	summon cats	*kɔ̀kórɛ̀*	summon chickens
mínū	summon cats	*mbórì*	summon goats
mûf	chase cats and dogs away	*mbórì ìràkà nàŋgà*	chase goats
mùndɛ̂	summon pigs	*míjǎw*	summon cats

DOI 10.15460/auue.2025.98.1.347

Mokpe		**Oroko**	
Form	**Function**	**Form**	**Function**
ndê	chase away dogs	*mínūs*	summon cats
ò lèmí	training dogs to not poo at the wrong place	*mûf*	cause pig to move
ōōōō wêj	summon dogs during hunting	*mŭúû*	summon cows
pŭs	summon cats	*ɲàkà*	summon cows
ʃ:	chase away chickens, ducks, goats	*nàŋgâ*	request dogs to lie down
ʃû	chase away goats, cats, dogs	*ndê*	summon goats
tɛ́mɛ̂	request dog, goat, pig to stand or stop moving	*ndʒákátù*	summon chickens
tìmbâ	chase away dogs	*ómá éêê ómâ*	incite dogs to chase game during hunting
úúw è	incite dog to chase game	*óŋgò*	summon dogs
wúʒrá-(nù)	chase away goats	*pŭs*	summon cats
‖-‖-‖	summon chickens, ducks	*ʃ:*	chase away chickens, goats, ducks
⊙-⊙-⊙	summon dogs	*ʃû*	chase away chickens
{clap-1}	summon pigs	*tókówâ*	request dogs to stand
{kiss-1}	summon dogs, pigs	*tʃâj*	chase away dogs, pigs, ducks
{object-1}	summon cats, dogs, chickens	*tʃàtʃàtʃàj*	chase away strange dogs
{object-2}	chase away cats, dogs, chickens	*‖-‖-‖*	summon chickens, cows

DOI 10.15460/auue.2025.98.1.347

Mokpe		**Oroko**	
Form	**Function**	**Form**	**Function**
{snap-1}	summon cats, dogs	*ǂ-ǂ-ǂ*	summon cow
{stamp-1}	chase away dogs	*⊙-⊙-⊙*	summon chickens
{tune-1}	incite dogs to chase game during hunting	*k'-k'-k'*	summon chickens, ducks
{whistle-1}	summon dogs, cats, duck	{kiss-1}	summon goats, chickens, cows
{whistle-2}	summon dogs	{object-1}	summon chickens
		{object-2}	chase away chickens
		{snap-1}	summon cats, dogs
		{whistle-1}	summon dogs, chickens
		{whistle-2}	summon dogs
		{whistle-3}	summon dogs

Below, we offer an analysis of the data provided in Table 1. First, in section 3.1, we describe the lexico-grammatical profiles of the CAC constructions in Mokpe and Oroko. In doing so, we will be guided by the prototypical features distinguished in scholarship (see section 2 above) grouped into two clusters: function-related properties (ecolinguistic and semantic features) and form-related properties (phonetic and morphological features).[7] Subsequently, we discuss the cognancy of Mokpe and Oroko CACs or the (more or less) pronounced lack thereof.

3.1 The functional and formal profile of CACs in Mokpe and Oroko

The lexico-grammatical profile of Mokpe and Oroko CACs pertains to the function (3.1) and form (3.2) of these constructions. In the

7 In this study, we will not discuss syntactic properties of CACs. The only syntactic feature that we tested – as part of the operational definition of CACs which we had used to include (or exclude) a construction into our dataset – is holophrasticity, i.e., the ability to function as an autonomous utterance.

DOI 10.15460/auue.2025.98.1.347

present section, we offer a detailed analysis of such functional and formal characteristics.

3.1.1 Function

The functional profile of Mokpe and Oroko CACs comprises of ecolinguistic (3.1.1.1) and semantic (3.1.1.2) properties.

3.1.1.1 Ecolinguistics

Speakers of Mokpe and Oroko are exposed to very similar climatic and environmental conditions. The Southwest region of Cameroon, where the two communities are located, belongs to an equatorial (monsoon) climate zone within Köppen climate classification. This climate type comprises of two major seasons: rainy (humid) and dry. The rainy season is characterized by a drop in temperatures and heavy downpours during July, August and September, which render the circulation of people and transport of goods difficult, especially in rural areas where the Oroko reside. Mokpe and Oroko settlements are surrounded by thick tropical forests. This creates some opportunity for hunting and, in some places, farming. Small streams in and around Oroko villages also avail the local population of micro-fishing activities for subsistence purposes. The Mokpe, many of whom live at the seacoast, practice fishing at a considerably larger scale.

Living in similar climatological conditions and being surrounded by similar flora, the Mokpe and Oroko also interact with similar animals. As far as wild animals are concerned, this includes elephants, antelopes, eagles, monkeys, rabbits, deer, bats, lizards, and snakes – just to name the species that are currently the most common and/or relevant for humans. With regard to domestic animals, the most widespread are goats, chickens, ducks, and pigs, as well as dogs and cats which are kept for hunting, defense, or protection-related purposes rather than as pets. While all the above-mentioned domestic species have their respective CACs in Mokpe and Oroko (see further below), no token directed to wild animals features in our database. This fact may stem from a limited range of hunting practices generally undertaken by the Oroko and, especially, more urban Mokpe communities, as well as the character of our fieldwork which, as explained above, did not involve hunters.

Contrary to several communities living in Cameroon and neighboring countries, Mokpe and Oroko people did not rear cows and

DOI 10.15460/auue.2025.98.1.347

sheep. Currently, cows and, to some extent, sheep may be found in Oroko settlements due to cultural contact with the Fulani. As a result, some cow-specific CACs have been developed and those that are directed to goats may also be occasionally extended to sheep. In contrast, given that the Mokpe do not live in an area adjacent to the Fulani, cows and sheep continue to be absent in Mokpe localities and CACs that would (specifically) be directed to cows and sheep are absent in the Mokpe data we collected. Certainly, during festive periods, the Mokpe interact with Fulani grazers who pass through Mokpe communities with their cattle for business purposes. The Mokpe may even purchase cows to slaughter for meat, but they do not rear them in their households.

In both Mokpe and Oroko communities, domestic animals, especially dogs, cats, and goats, are given proper names. Four types of animal names are attested. First, an animal can be named after deceased loved ones in remembrance of them. For example, the name *Èjóndɛ̀* in Oroko was given to a cat in memory of a deceased grandmother who bore the same name (see also the dog name *Màŋákɛ̂* lit. 'wonderful' in Mokpe). Second, domestic animals can be named after certain wild species. For instance, dogs may be given the proper name *Lájɔ̀n* in Mokpe borrowed from English 'lion'. Third, a name may reflect the location from where the animal originated. This can be illustrated with the Mokpe name *Wɔ̀ndʒɔ̀* which was given to a goat that had been brought from the village of *Wɔ̀ndʒɔ̀ŋgɔ̀* located in Fako division. Fourth, the name of an animal can refer to a character trait or physical appearance. Such names are mostly borrowed from English, e.g. *Ràtì frɔ́gì* – a name that is typically given to cats because of their uncanny ability to hunt rats and frogs – or *Rambo*, *Bullet*, *Thunder*, and *Smart* which suggest the perceived strength, speed, and intelligence of a dog.

3.1.1.2 Semantics

Nearly all the CACs included in our database convey actions related to motion. This is the rule in Oroko where all tokens indeed express an idea of motion. In Mokpe, one CAC, i.e. *ò lèmí* does not principally imply motion but is rather used to shame dogs for discharging feces in inappropriate places and to train them not to do so. Nevertheless, even this CAC may trigger the animal to go away and may thus be related to motion at least minimally. (It should be noted that *Ò lèmí*

DOI 10.15460/auue.2025.98.1.347

is a secondary CAC which literally means 'Are you crazy?' The lexicalization of *ò lèmí* into a CAC (or its CAC-ization) is low and the construction is widely used with human addressees as well.) The three semantic motion-related categories of CACs found across languages, i.e. summonses, dispersals and directionals, are attested in both Mokpe and Oroko. In the two languages, summonses are the most common types of CACs; dispersals are the second most common; and directionals are the least common. Accordingly, the semantic hierarchy of actions exhibited by Mokpe and Oroko CACs coincides with those identified for two other Western Cameroonian languages, i.e. Babanki and Bum (Andrason & Akumbu 2024).

With regard to the distribution of CACs among domestic species – which, as we indicated above, are the only addressees of CACs in Mokpe and Oroko – the following can be observed: In Mokpe, the largest group of CACs is used with dogs and, slightly less so with cats. The next class of recipients of CACs includes chickens, goats, and pigs. Each of these species has approximately half of the CACs associated with dogs or cats. Ducks have the fewest CACs. Oroko data reveal a similar, although not identical, hierarchy. Dogs and, somewhat less, chickens can be addressed with the largest number of CACs. The number of CACs compatible with goats and cats is approximately half this number. Pigs and cows have yet fewer CACs associated with them, and ducks, as in Mokpe are the least frequent addressees of CACs.

Primary CACs are uniquely directed to animals and their usage with human speakers would be perceived as extremely offensive. In contrast, secondary CACs may be employed in communication with both animals and humans and generally do not trigger pejorative effects (other than their literal meaning; see *ò lèmí* discussed above).

Overall, Mokpe and Oroko CACs, especially primary ones, are often monosemic: they express a specific action and are directed to a specific animal species. Many CACs that are limited to a particular species are summons, some of them exhibiting an onomatopoeic foundation, e.g. *kùrúkùrú* (restricted to chickens) and *mɛ̌ɛ́ɛ̀?* (restricted to goats) in Mokpe and *mŭúúû* (restricted to cows) in Oroko. However, several other CACs may be used with a range of species, especially if such animals are phenotypically similar (e.g. small). For instance, in Mokpe, the primary CAC *ʃ:* can be employed with chickens, ducks, and goats, while *ʃû* can be employed with cats, dogs, and goats. Sim-

DOI 10.15460/auue.2025.98.1.347

ilarly, the secondary CAC *àŋdʒíjá áŋgâ* meaning 'leave that place' is used to chase away pigs, dogs, and cats and the borrowed *mûf* can be used to chase away both dogs and cats. The same phenomenon is observed in Oroko. The primary CAC ‖ may be addressed to both chickens and cows and the secondary CAC *árá wàgíjâ* 'leave that place' to chickens and cats. The polysemous character of secondary CACs is significantly more pronounced than is the case of primary CACs and it is likely that many secondary CACs can be used with a larger set of animals than attested in our data.

3.1.2 Form

The formal profile of Mokpe and Oroko CACs comprises of phonetic (3.1.2.1) and morphological (3.1.2.2) properties.

3.1.2.1 Phonetics

Mokpe and Oroko CACs exhibit distinct phonetic profiles depending on whether the specific CAC forms are primary or secondary. In general terms: primary CACs tend to be short, more consonantal than vocalic, and, to some extent, extra-systematic, whereas secondary CACs are robust, equally vocalic and consonantal, and systematic.

Regarding phonetic shortness/robustness, monosyllabicity is typical of primary CACs in Mokpe and Oroko, while for secondary CACs, polysyllabicity is far more common. To be exact, the vast majority of primary Mokpe/Oroko CACs are monosyllabic (see *hrɔ̀ɔ̀* and *mɛ́ɛ̀ʔ* in Mokpe and *bɛ́ɛ̀* and *fííí* in Oroko) or draw on identical monosyllabic segments arranged into sequences (see *hôhôhô* in Mokpe and *kɔ̀kɔ̀kɔ̀* in Oroko; see next section). In contrast, very few primary CACs are disyllabic (e.g. *míjǎw* in Mokpe and Oroko) or constitute a series of identical dissyllabic segments (e.g. *kùrúkùrú* in Mokpe). For secondary CACs, this tendency is reversed. The vast majority of secondary forms are polysyllabic: disyllabic (e.g. *dʒànâ* 'bring (it)' and *lèmbɛ̂* 'catch (it)' in Mokpe and *jáká* 'come' and *kómá* 'chase/follow' in Oroko) or trisyllabic (e.g. *ámbɛ́lɛ̂* 'wait' and *àŋdʒíjá* 'leave' in Mokpe and *tókówâ* 'stand' and *àráwâ* 'go/leave' in Oroko). If all the constituents of secondary CACs that originate in analytical phrases are considered jointly, syllabic length may even be greater (see *érákâ ùndáwùnɛ̂* which literally means 'leave the house' and contains 7 syllables). By contrast, monosyllabic secondary CACs are extremely rare (see *ndâ* 'go/leave' in Mokpe/Oroko and *gbâ* 'follow' in Orko). The tendency of

DOI 10.15460/auue.2025.98.1.347

secondary CACs to consist of two syllables is overall consistent with the prevalence of disyllabic roots in the general word stock of Mokpe (Atindogbe 2013) and Oroko (D. Friesen 2002).

In Mokpe and Oroko, CACs tend to draw on both consonants and vowels. This relatively equal contribution of consonantal and vocalic material to the phonetic substance of CACs is particularly evident in secondary CACs. Indeed, in our data, there are no secondary CACs that would only consist of vowels or, on the contrary, consonants. Nevertheless, as far as primary CACs are concerned, their consonantal nature seems somewhat more marked. First, in both languages, a number of primary CACs are exclusively built around consonants as illustrated by *ʃː*, *ǁ-ǁ-ǁ*, and *ʘ-ʘ-ʘ* in Mokpe and Oroko, and *k'-k'-k'* and *ǂ-ǂ-ǂ* in Oroko. In contrast, purely vocalic primary CACs are rare and attested only in Mokpe: *ōōōō* in *ōōōō wêj wêj* (see also *úúw è hèhèhèhè*). Second, except for *ōōōō wêj*, and *úúw è hèhèhèhè* mentioned above, primary CACs regularly exhibit consonantal onsets – a fact that clashes with a much larger number of onset-less secondary CACs (e.g. *ámbélɛ̂*, *àŋdʒíjá áŋgâ*, *àráwâ*, *èràkà*, *ómá éɛ̂é ómâ*, and *óŋgò*). Third, primary CACs (e.g. *míjǎw* in both languages, *pǔs* in Mokpe, and *tʃâj* in Oroko) allow for consonantal codas – a phenomenon that is absent in secondary CACs and generally disallowed in the standard phonetic systems of the two languages.[8]

The majority of CACs are wholly built around systematic phonetic material, i.e. consonants and vowels that form part of the standard phonetic repertoire of Mokpe and Oroko. While this phonetic systematicity prevails, some CACs – invariably primary ones – make use of extra-systematic sounds or even are solely made up of extra-systematic phones.

One class of the extra-systematic sounds found in Mokpe/Oroko CACs involves phones that, although absent in the standard phonetic repertoires of these two languages (see Neh 1989 and Atindogbe 2013 for Mokpe and Atta 1993 and D. Friesen 2002 for Oroko), are attested in the phonetic systems of other languages and feature in the International Phonetic Alphabet. In total, there are five such extra-systematic IPA sounds in the collected data. Three of them are clicks: the lateral [ǁ] and the bilabial [ʘ] (attested in both Mokpe and Oroko), as well as the palatal [ǂ] (attested only in Oroko) – each

8 This phenomenon is also attested in CAC loanwords (see further below).

DOI 10.15460/auue.2025.98.1.347

of which usually appears in a series. Neither of the above-mentioned clicks combines with other phones, whether systematic or extra-systematic. The remaining extra-systematic consonants attested in CACs are the velar ejective [k'] and the voiceless postalveolar fricative [ʃ], which are found in CACs but are absent in the regular word stock in Oroko (D. Friesen 2002); the trill [r] and the labio-dental fricative [f], which except for CACs, do not feature in Mokpe; and the glottal fricative [h] which in both Oroko and Mokpe is limited to CACs and interjections. The extra-systematicity of the vocalic material used in CACs is much less pronounced. Our data do not reveal vowels of extra-systematic quality. Nevertheless, in both Mokpe and Oroko, vowels found in CACs can be nasalized (see Mokpe *hrɔ̃ɔ̃̃* and Oroko *hrɔ̃̃*) contrary to the standard phonetic systems of these languages that only consist of oral vowels.

The other, equally considerable, set of extra-systematic sounds found in Mokpe/Oroko CACs comprises articulations that are not included in the International Phonetic Alphabet. Such non-IPA sounds may be produced orally, i.e. with the vocal tract, or non-orally, i.e. without the use of a vocal-tract mechanism. Non-IPA oral sounds attested in Mokpe and Oroko CACs are kisses, whistles, and the so-called tune. As was the case of clicks, these sounds never combine with other phones in CACs. {Whistle-1}, attested in both Mokpe and Oroko, is a sequence of short high-tone and high-pitch whistles. Following the notation which we developed in our previous study on CACs in Babanki and Bum (see Andrason & Akumbu 2024) – which reflects the proposal originally put forward by Poyatos (1993, 2002) – we represent this whistle as [S˥-S˥-S˥(…)]. {Whistle-2}, also attested in both languages, exhibits a more complex tonal structure: it begins with a rise in pitch and ends with a pitch fall, i.e. [S˥˩]. Lastly, {whistle-3}, which is restricted to Oroko, is a series of rising-pitch short whistles, i.e. [S↗-S↗-S↗(…)]. {Kiss-1} is found in both Mokpe and Oroko. It matches a cross-linguistically common kiss-sound found in CACs that draws on a pressure-release mechanism produced by two closures: an anterior dorsovelar one (made with the tongue) and a forward labial one (made with protruded lips) (Andrason 2024). In scholarly literature, this speech-kiss sound has been represented with the tripartite symbol [↓B'] that refers to the ingressive [↓], labial [B], and click-like nature ['] of this phone (Poyatos 1993, 2002, Andrason & Karani 2021, Andrason 2024). Similar to clicks, {kiss-1} often

DOI 10.15460/auue.2025.98.1.347

appears in a series. The remaining non-IPA sound produced orally is a melodic vocalization found in Mokpe which we label as {tune-1}. This repeated song-like pattern uses high-toned back high and front mid vowel patterns followed by a verb that encourages or motivates a dog during hunting.

The remaining types of non-IPA extra-systematic sounds are made with body parts other than the mouth and vocal tract, namely fingers (snaps), hands (claps), and feet (stamps), as well as with objects. To be exact, {snap-1}, found in Mokpe and Oroko, is a short finger snap usually realized in a series. In like manner, {clap-1} found in Mokpe is a serial realization of short claps done by slapping the two hands against each other. {stamp-1} is produced by periodically hitting one foot against the ground to chase away animals. In addition, speakers may use objects to communicate with animals. The CAC referred to as {object-1} is made by means of a dish that usually contains food given to an animal. The sound is produced by repeatedly hitting the dish with another object, against a wall, or on the ground. Whatever object is used to hit the dish, the sound produced serves the same function of inviting animals to eat. {object-2} is a louder sound typically made with a piece of wood (however, the use of non-wooden tools is also possible). It may be realized as a single "beat" or a series of "beats" as is necessary to catch the attention of the animal.

As far as phonotactics are concerned, primary and secondary CACs also exhibit different properties. Secondary CACs make use of syllable structures that are fully systematic in both Mokpe and Oroko (see Neh 1989 and Atindogbe 2013 for Mokpe and Atta 1993 and D. Friesen 2002 for Oroko), namely CV (e.g. *βέβέlê* and *tókówâ*) and V (see the first syllable in *àráwâ* and *ómá*), as well as, with nasals only, NCV (see *mbórì/mbólì*) or CVC (*kám*). In contrast, primary CACs may additionally exhibit a C syllabic structure, thus allowing for consonantal nuclei and non-vocalic words. The exemplary cases are *ʃː*, *ǁ-ǁ-ǁ*, *⊙-⊙-⊙* in Mokpe and Oroko, as well as *ǂ-ǂ-ǂ-ǂ* and *k'-k'-k'* in Oroko. This type of syllable (and word) structure is unattested in the general word stock of Mokpe and Oroko. A further phonotactic peculiarity of primary CACs concerns onsets and codas. Both in Mokpe and Oroko, CACs tolerate complex onsets in which C1 is a consonant other than a nasal; see [#hr-] in Mokpe *hr̃ɔ̃ɔ̃* and Oroko *hr̃ɔ̃*. CACs also allow for non-nasal codas, e.g. approximants (*míjǎw* in Mokpe and *tʃâj* in Oroko) and genuine consonants, specifically, [s] (*pǔs* in both lan-

DOI 10.15460/auue.2025.98.1.347

guages). CACs that are borrowed exhibit another extra-systematic coda, namely [-f#] (see *mûf* 'move' from English into both Mokpe and Oroko).

Lastly, primary CACs distinguish themselves from the general word stock, as well as from secondary CACs, by exhibiting two or three degrees of length. As far as we know, vocalic length is not a contrastive feature in the general word stock of Mokpe and Oroko even if some vowels can be lengthened (see Neh 1989: 66–67 and Atindogbe 2013 for Mokpe and Atta 1993: 60 and D. Friesen 2002 for Oroko). In primary CACs, long vowels are not only attested but also relatively common and 'stable', as illustrated by Mokpe *hr̃̃ɔ̃̃*, *mɛ́ɛ́ɛ̀ʔ*, and *ōōō wêj* and Oroko *bɛ́ɛ́ɛ̀*, *fííí*, and *mŭúûû*. That is, the above-mentioned CACs are not mere lengthened variants, but rather the regular forms that are typically realized with well-audible long vowels. The most frequent length type in such CACs is an exaggerated or extra-long realization, equivalent to three morae (see Mokpe *mɛ́ɛ́ɛ̀ʔ* and Oroko *bɛ́ɛ́ɛ̀*). Nevertheless, 'non-exaggerated' long vowels (equivalent to two morae) are also found (see *hr̃̃ɔ̃̃* in Mokpe). Consonants may also be long in primary CACs in Mokpe and Oroko. The exemplary case is *ʃ̩ː*.

3.1.2.2 Morphology

Similar to phonetics, the morphological profiles of primary and secondary CACs are quite dissimilar in Mokpe and Oroko. Primary CACs are morphologically simple, whereas secondary CACs are (or at least may be) complex.

All primary CACs in Mokpe and Oroko are monomorphemic. They consist of a root – the so-called CAC matrix – and do not contain any other morphemes whether inflectional, derivational, or compounded. This morphological simplicity is only violated in cases where a CAC is a series of identical segments, e.g. *ǁ-ǁ-ǁ*, *⊙-⊙-⊙*, *tʃàtʃàtʃàj*, *ǂ-ǂ-ǂ-ǂ*, *k'-k'-k'*, and *kùrúkùrú*. In agreement with recent work on CACs, we regard such replications as expressive/phonetic rather than derivative/morphological strategies. As is true of CACs in many other languages, this view can be substantiated by the following facts. First, the use of isolated segments (or singletons) outside of serialized CACs is problematic. Second, the addition of another segment and, for example, the expansion of a triplicated structure to quadruplicated, does not trigger a change in meaning. The action requested

DOI 10.15460/auue.2025.98.1.347

and the animal concerned remains the same. Third, replicated CACs need not have an intensifying effect, as illustrated by triplicated ‖-‖-‖ and ⊙-⊙-⊙ which are not more intense or emphatic than duplicated or non-replicated summonses (e.g. *kùrúkùrú*). This same applies to *ʧàtʃàtʃàj* which is used in Oroko to chase away unknown dogs and comprises three *ʧà*(*j*) chunks. To be sure, Oroko contains another dispersal CAC, i.e. the singleton *ʧâj* used to chase away dogs, pigs, and ducks (furthermore, the same form *ʧâj* is used in Cameroonian Pidgin English (CPE) as an exclamation to express surprise in moments of sadness). Nevertheless, the CAC *ʧàtʃàtʃàj* need not be viewed as tri-morphemic with each *ʧâj* segment conveying a separate meaning. Rather, *ʧàtʃàtʃàj* constitutes a holistic pattern indivisible into more elementary meaning-bearing units – despite its relation to the CAC *ʧâj* and, much less likely, the (emotive) interjection *ʧâj*.

Contrary to primary CACs, secondary CACs can – although need not – be morphologically complex. In cases where this complexity is attested, it reflects the non-CAC source underlying a given CAC construction. It thus stems from the fact that such CACs originate in other lexical classes or phrases/clauses built around non-CAC constituents. Most secondary Mokpe/Oroko CACs derive from imperatives, with which they are still homophonous, e.g. *ámbɛ́lɛ̂* ‘wait!’, *βɛ́βɛ́lɛ̂* ‘chase!’, *ʤànâ* ‘bring!’ in Mokpe and *àráwâ* ‘go!’, *kómá* ‘chase!’, and *tókówâ* ‘stand!’ in Oroko. Given their imperative origin, these CACs exhibit imperative morphology. In Mokpe and Oroko, imperatives are (one of) the simplest forms of a verb, although they are marked by their own, very diverse tonal patterns (Atindogbe 2013: 82). As a result, it is possible to argue that, similar to the corresponding imperatives, (de)imperative CACs exhibit at least a bi-morphemic structure: they consist of a root (and other extensions, should this be the case) and an imperative tone configuration. Comparably, CACs derived from nouns – *mbólì/mbórì* ‘goat’ in Mokpe and Oroko and *ɲàkà* ‘cow’ in Oroko – may be regarded as bi-morphemic (or even more complex). For example, in agreement with their nominal source, the CAC *mbólì/mbórì* consists of the root *bólì* and the class-9 nasal prefix (in this case, *m*-) and *ɲàkà* ‘cow’ consists of the root *-ɲàkà* and a zero class-9 prefix (since the root begins with a nasal, see Atindogbe 2013: 32). The morphological complexity of secondary CACs is even greater if a CAC construction descends from a small phrase or clause, e.g. one built around an imperative and a noun (Oroko *gbɛ́ ò ndáwò* ‘enter the

DOI 10.15460/auue.2025.98.1.347

house'), or a series of imperatives (Oroko *ìràkà nàŋgà* 'go and sleep)'. It should however be noted that the lexicalization of such original analytical phrases in their CAC uses is low.

CACs that are borrowed from other languages may also reflect forms that are morphologically complex in the donor language system. For instance, Mokpe *gō áùt* is a CPE phrasal verb *go out*, composed of the verbal root (imperative) and a particle; *kám yéè* 'come here' is a CPE expression composed of an imperative and a locative adverbial; and *líf dá plēs* is an adaptation of the small clause 'leave that place' composed of three free morphemes. This is also the case of Oroko *kátʃàm* 'catch it' which derives from a bi-morphemic construction *catcham* in CPE. However, some expressions (e.g. Mokpe *bâk* 'back' and *ínsáì* 'inside') reflect monomorphemic forms in the donor language. Most borrowed CACs originate from Cameroon Pidgin English, with Mokpe containing a larger number of CPE loans (5x) than Oroko (2x). This presence of CPE CACs might reflect the multilingualism of the interviewed speakers, who could switch between Mokpe/Oroko, CPE and, perhaps, even English when talking to animals. Thus, rather than borrowing *sensu stricto*, we would deal with some type of code switching. Nevertheless, since code switching and borrowing are connected phenomena constituting two edges of a shared continuum (Matras 2009, Pakendorf 2009, Meakins 2011, Gardani 2020, Andrason 2021), we regard such CACs as some types of loanwords, whether less general/entrenched/panlectal/integrated/default (more code-switching-like) or more general/entrenched/panlectal/integrated/default (more borrowing-like).

The only CAC that can be traced to French – one of the former colonial languages still used in Cameroon – is *mínū* in Mokpe and *mínūs* in Oroko. Both CACs are employed to summon cats as is also true of their French source *minou*. The presence of the final *-s* in the Oroko form is difficult to explain. It is possible that *s* found in the coda is analogical to *pŭs* which is found in both Mokpe and Oroko and entertains the same function: a summons to cats.[9] The borrowing of CACs from both Cameroon Pidgin English and French on the one

9 Compare with the CAC *les* [lés] in Ewe. *Les* consists of the Ewe imperative verb *le* 'catch!' and a suffix *-s*. The suffixation of *-s* is attributed to analogy with other CACs found in Ewe and "the 'naturalness' or the crosslinguistic pressure of closing mono-syllabic CACs with consonantal codas and using sibilants for these purposes" (Andrason & Gafatsi 2025: 19).

DOI 10.15460/auue.2025.98.1.347

hand and the larger number of CPE loans than French loans on the other, stem from the fact that, as explained in the introduction, Mokpe and Oroko are spoken in the Anglophone zone of Cameroon (see D. Friesen 2002). The greater visibility of CPE loans in Mokpe is, in turn, likely due to the extensive presence of Cameroon Pidgin English in the area which includes parts of Cameroon's South-West Regional headquarters and, is therefore, more urbanized than the area where Oroko is spoken. Borrowed CACs tend to be directionals, slightly less often dispersals, and only residually summonses. Accordingly, they contrast with primary CACs for which summonses are more common than dispersals and directionals.

To conclude, our data demonstrate that, when considered in their totality, the categories of CACs in Mokpe and Oroko are opaque. There are no morphological patterns that would be pervasive in and/or exclusive to CACs. On the contrary, CACs allow for a great diversity of word structures ranging from strictly monomorphemic (and simple) to pluri-morphemic (and complex).

3.2 Relatedness

As was the case with the phonetic and morphological properties of Mokpe/Oroko CACs, which are largely conditioned by the primary or secondary status of CAC constructions, primary and secondary CACs exhibit different extents of similarity. In general terms, primary CACs tend to be more similar in the two languages than secondary CACs. Among the primary CACs attested, 8 are identical or highly similar in both languages. For Mokpe, this amounts to slightly more than half of all primary CACs (i.e. 8 out of 14 = 57%); for Oroko, this number constitutes slightly less than half (i.e. 8 out of 17 = 47%). The following CACs are identical in both form and meaning in Mokpe and Oroko: *hrɔ̃ɔ̃*, *kɔ̀kɔ̀kɔ̀*, *míjǎw*, *pŭs*, *ʃː*, and *ʃû*. The majority are summonses (4x). The remaining type comprises of dispersals (2x). No directionals or motion-unrelated CACs are identical in both languages. The cognancy of the Mokpe/Oroko general word stock – or more correctly, the 40 words that we collected following the Automated-Similarity-Judgment-Program (ASJP) list (Wichmann 2007, Wichmann, Holman & Brown 2022) – ranges between 65% and 75%.[10] The following 23

10 The ASJP list (Wichmann et al. 2012, Wichmann, Holman & Brown 2022) can be found in the Max Planck repository. Initially, the list drew on the Swadesh list and thus included 100 lexemes (Brown et al. 2008). However, "research shows that

DOI 10.15460/auue.2025.98.1.347

(Mokpe/Oroko) ASJP lexemes may be regarded as cognate: EYE (*lízrɔ̀/lísɔ̀*); EAR (*lìtô/lìtó*); TONGUE (*dʒémɛ̀/ìjémé*); TOOTH (*lìʒrɔ̀ŋgá/lìsɔ̀ŋgá*); KNEE (*lùwɔ́ŋgɔ́ŋgɔ́/lìβɔ́ŋgɔ̂*); BLOOD (*màjdʒá/màkìjá*); BONE (*èèʒré/èsé*); BREAST (*lìwɛ̂/màβɛ́*); LIVER (*lìwàj/lìβɛ̀*); LOUSE (*ɲìɲá/ɲa, nija, ɲia* (D. Friesen 2002)); DOG (*ŋgbâ/ŋgbá*); FISH (*ɲàmà màlúwá/ɲamamaliba* (D. Friesen 2002)); HORN – ANIMAL PART (*mòʒréwá/mòséβâ*); PERSON (*màlúwá/ìlíβâ*); NAME (*ndʒìjá/ndʒìjá*); DIE (*w̃â/wa* (D. Friesen 2002)); SEE (*énɛ̂/énɛ̂*); ONE (*βɔ́kɔ́, jɔ̀kɔ́/èjɔ́kɔ́*); TWO (*wéwă/béwɛ̄*); I (*ìmbâ/mbá*); YOU (*òwâ/ùwá*). The above shows that, for Oroko, the cognancy of CACs is considerably lower than that of the ASJP vocabulary; for Mokpe, it is also lower, although less radically.

Importantly, out of the four summonses that are identical in Mokpe and Oroko, three have an onomatopoeic foundation. Accordingly, they coincide with onomatopoeias (also attested in Mokpe and Oroko) that imitate the sounds made by the respective animals: *hrɔ̃̀ɔ̃̀* by pigs, *kɔ̀kɔ̀kɔ̀* by chickens, and *míjằw* by cats. This imitative nature is also patent through the similarity of these Mokpe/Oroko forms with onomatopoeias attested in many other languages, including those that are typologically, geographically, and phylogenetically remote and/or disconnected. For example, *hrɔ̃̀ɔ̃̀* has close onomatopoeic equivalents in several Slavonic languages, such as Czech (*chro*), Polish (*chrum*), and Russian (*хрю* [xr'ju]), as well as Romanian (*groh*) and Indonesian (*grok*). This cross-linguistic correspondence is even more evident with the two other tokens. CACs comparable to *míjằw* are, for example, found in Kihunde (*mìjáw*(*ù*)), Asante, Bono, and Fante (*mìàẁ*), Macha and Harar Oromo (*miw* and *mew*, respectively), with onomatopoeias exhibiting forms similar to *míjằw* being even more prevalent cross-linguistically.[11] CACs comparable to *kɔ̀kɔ̀kɔ̀* are attested in Arusa Maasai (*kúkukúku*), Fante (*kúkú*(*kú*)), and Sengwer (*kutkutkutkut*). Again, onomatopoeias of this type are even more common across languages, e.g. in Bulgarian (*ко-ко-ко*), Greek (*ко-ко*), Japanese (*kokko*), Portuguese (*có-có-có*), and Tjwao (*kokokokoko*). The other identical CAC in Mokpe and Oroko, i.e. the summons *pŭs*

automated language classification is just as accurate, if not more so, using a much shorter list restricted to the 40 most stable items of the 100-item list" (Wichmann 2007: 1, Wichmann, Holman & Brown 2022).

11 Additionally, in many languages, the nasal onset [ɲ] appears instead of [m], e.g. *ɲá:u* in Arusa Maasai.

DOI 10.15460/auue.2025.98.1.347

directed to cats, is also frequent from a crosslinguistic perspective. Indeed, similar CACs addressed to cats are found in Dutch (*poes*); with a front close vowel instead of the back one, they are found in Romanian (*pis*(*i*)), Turkish (*pisi*), and Azeri (*pish*); with no vocalic nucleus, in Ayt-Haddidu Tamazight and Sengwer (*ps*); and with the voice bilabial stop instead of the voiceless one, in Arabic (*bis*), Babanki (*bùús*), and Bono (*bú(ú)s*). Although the CAC *pŭs* found in Mokpe and Oroko and similar forms attested in other languages need not have an onomatopoeic foundation imitating meowing, they may mimic a more aggressive sound produced by cats when hissing (cf. Schötz 2020).[12]

As is the case of the summonses described above, the two dispersals that exhibit the same form in Mokpe and Oroko, i.e. *ʃː* and *ʃû*, exploit cross-linguistic strategies. In the sample of 79 languages (Andrason 2023), dispersals tend to draw on sibilants, with [ʃ] being the most commonly attested. A non-vocalic syllable pattern similar to *ʃː*, i.e. [Ṣ] (that is, a syllabic sibilant), appears in a tenth of sibilant dispersals, while the SV syllable structure, as in *ʃû*, is slightly more frequent. The most common vowels are close, specifically U- and I-types – each featuring in approximately a third of the language sample underlying that study. For example, a dispersal form *ʃ* is found in Arusa Maasai, Dza, German, Lithuanian, Malay, Matses, Nahuatl, Persian, and Polish. A dispersal form *ʃU* is found in Malay, English, Ndebele, Oromo, Shona, Suba, and Xhosa. While not necessarily onomatopoeic, the presence of sibilants in dispersals is motivated (Andrason 2023).

The crosslinguistic pervasiveness – whether due to an onomatopoeic foundation or motivated by other reasons – means that although identical, the above-mentioned CACs in Mokpe and Oroko need not have descended from a common ancestor. Equally likely (perhaps, in fact, more so), they may have emerged by harnessing common cross-linguistic strategies, independently in Mokpe and Oroko. Such parallel developments are even more plausible in light of the two other Mokpe/Oroko CACs that have the same form, which we discuss in the next paragraph.

12 The majority of the examples of CACs from languages other than Mokpe and Oroko provided in this paragraph are extracted from *A living database of conative animal calls. Version 1.1* (Andrason et al. 2024). A few others come from a forthcoming edition of this database, i.e., version 1.2.

DOI 10.15460/auue.2025.98.1.347

Two primary CACs coincide formally in Mokpe and Oroko but differ as far as their functions are concerned. Both are clicks, i.e. *ǁ-ǁ-ǁ* and *⊙-⊙-⊙*. In Mokpe and Oroko, *ǁ-ǁ-ǁ* is used to summon chickens. However, the same CAC can also be addressed to other species: in Mokpe to ducks while in Oroko to cows. Across languages, the lateral click *ǁ* – employed typically on its own with no accompanying phones whether consonants or vowels – is widely exploited in CACs and entertains a variety of functions. For example, to summon animals (Arusa, Ewe, Macha Oromo, Babanki, and Bum), modify the motion of an animal (Arusa, Tjwao, Ewe), chase away an animal (Tjwao), and request actions that are not related to motion, such as silencing (Tjwao), encouraging animals to give milk (Sengwer), and merely drawing the attention of an animal (Ewe). The click *ǁ* is also compatible with a number of different animal species: goats, sheep, donkeys, cows (cattle), dogs, and poultry. Like *ǁ-ǁ-ǁ*, *⊙-⊙-⊙* is employed to summon animals in Mokpe and Oroko. However, in Mokpe, it is directed to dogs, while in Oroko to chickens. The use of *⊙* in a summoning function is widely attested across languages: with dogs (Arusa and Sengwer) and specifically puppies (Bono), livestock (Tjwao), poultry (Kihunde and Asante), goats and sheep (Asante). Additionally, although much less common, *⊙* can be employed as a directional to modify the motion of animals (in Fante). Overall, click CACs (including the post-alveolar [!] and the palatal [ǂ]) are highly pervasive from a typological perspective. Indeed, nearly all languages that have been studied thus far have at least one click CAC in their repertoire (Andrason 2024). As far as non-click languages are concerned, click consonants constitute the most ubiquitous class of extra-systematic IPA sounds found in CACs (Andrason & Karani 2021: 18, Andrason 2024). Therefore, it is likely that the presence of the same click CACs in Mokpe and Oroko reflects a crosslinguistic pressure rather than stems from shared ancestry (although the latter cannot be ruled out either). The different meanings of these forms would be a palpable manifestation of this.

The remaining primary CACs are distinct in Mokpe and Oroko. There are 5 CACs in Mokpe that have no equivalents in Oroko: *hôhôhô*, *kùrúkùrú*, *mɛ́ɛ̀ʔ*, *ōōōō wêj*, and *úúw è hèhèhèhè*. In Oroko, nine CACs do not have equivalents in Mokpe: *bɛ́ɛ̀*, *fíí́*, *kǎʃà*, *kɔ̀kɔ́rɛ̀*, *mǔúû*, *tʃâj*, *tʃàtʃàtʃàj*, *k'-k'-k'*, and *ǂ-ǂ-ǂ-ǂ*. The majority of such dissimilar primary CACs are summonses several of which have an onomatopoeic foun-

DOI 10.15460/auue.2025.98.1.347

dation, e.g. *kùrúkùrú*, *mɛ̌ɛ̀ɛ̀ʔ*, *bɛ̌ɛ̀ɛ̀*, *kɔ̀kɔ́rɛ̀*, *mǔúûʉ*, and perhaps even *k'-k'-k'*. Once again, similar CACs (or onomatopoeias) are widely attested in the languages of the world, as illustrated by *kukuryku*, *mee*, *bee*, *muu* in Polish and other Slavonic languages (Siatkowska 1976). Therefore, the above-mentioned CACs have most likely been developed in Mokpe and Oroko by harnessing well-known crosslinguistic strategies – although independently. The lexemes *tʃâj* and *tʃàtʃàtʃàj* with which speakers chase away animals, as we explained in section 3.1.2.2, exploit "universal" mechanisms as well. They draw on sibilant affricates which are common in dispersals and the nucleic vowel *a* which is almost as common as the U- and I-type vowels mentioned above (and discussed in Andrason 2023). The palatal click *ǂ-ǂ-ǂ-ǂ*, which is found in Oroko but unattested in Mokpe, is also common in summonses in other languages. It is used to summon poultry and/or cats in Asante, Bono, Fante, and Arusa. Even *kǎʃà* employed to incite dogs to chase prey may be explained in terms of common typological tendencies. This CAC makes use of the plosive [k] and a sibilant [ʃ] and thus closely approximates the form exhibited by a prototypical dispersal [kI/Uʃ] (Andrason 2023), with the meaning of chasing prey (a directional subtype) and chasing away an animal (dispersal) being related. (This CAC may also be related to *kátʃàm* and thus reflect the CPE form *catcham*.)

Overall, drawing on common crosslinguistic strategies seems to constitute the chief manner of coining primary CACs in Mokpe and Oroko. Although this may lead to the emergence of forms that are similar (formally/functionally or formally only) in both languages – apart from the creation of CACs that are clearly different – to account for such similarities one need not involve proto-Bafawic-Bakweric origin. In other words, our analysis shows that none of the CACs that are currently similar in Mokpe and Oroko necessary descends from a shared ancestor. Parallel developments drawing on common crosslinguistic strategies are much more likely.

In contrast to what we observe with primary CACs, the similarity of secondary CACs in Mokpe and Oroko is very low. There are only two unquestionable cognates in our data. The first is the pair *mbólì* (Mokpe) and *mbórì* (Oroko). Both CACs derive from a natural-kind label, i.e. the noun meaning 'goat' and are used to summon goats (see section 3.1). The other is *ndê* which derives from a homophonous imperative verb that means 'go' (and thus 'leave') in Mokpe

DOI 10.15460/auue.2025.98.1.347

and ‘come’ in Oroko. As a CAC, *ndê* is respectively used as a dispersal (with dogs) and a summons (with goats). The remaining secondary CACs do not exbibit any similarity (although they do draw on verbal/nominal/adverbial roots that are cognate). Nevertheless, although independently, both Mokpe and Oroko exploit the same “universal” strategies that regulate the ‘cooptation’ of some lexical classes for human-to-animal directive purposes. Specifically, as is common across languages (and as we explained in section 3.1), imperatives constitute the most prolific source of secondary CACs, in Mokpe (e.g. *ámbélê* ‘leave’, *βéβélê* ‘chase’, *dʒànâ* ‘bring’, *kám* ‘come’, *lìjâ* ‘sit’, *témê* ‘stand’, *tìmbâ* ‘return’) and Oroko (e.g. *àráwâ* ‘go’, *èràkà* ‘go’, *gbâ* ‘kill’, *jákâ* ‘come’, *kómá* ‘chase’, *nàŋgâ* ‘sleep’, *óŋgò* ‘come’, and *tókówâ* ‘stand’). In some instances, a CAC is a small phrase in which the imperative is accompanied by a locative adverbial or nominal (e.g. *nù* ‘here’, *yéè* ‘here’, *wàgíjâ* ‘this place’, *ò ndáwò* ‘in the house’).[13] The other source of secondary CACs, significantly less frequent in Mokpe and Oroko (as is also the case crosslinguistically), are natural kind labels: *ɲàkà* ‘cow’ in Oroko and the above-mentioned *mbólì/mbórì* in the two languages.

As we also explained in section 3.1, a few CACs are borrowed from other languages. Among such loanwords, two are shared by both Mokpe and Oroko. The first of them is *mûf* adopted from the CPE imperative *move*. *Mûf* is used as a dispersal to chase away cats and dogs in Mokpe, but as a directional to make pigs advance in Oroko. Both meanings can be connected to the semantic content of their CPE imperative source. The other borrowed CAC that coincides in Mokpe and Oroko is *mínū/mínūs* used to call cats. As we explained above, this CAC is most likely adopted from the homophonous and synonymous summons *minou* found in French. The remaining CACs have been borrowed independently in Mokpe and Oroko. Mokpe contains at least four such CACs, while Oroko contains one CAC loanword, which is also absent in Mokpe.

13 As locative elements seem optional and may be changed or omitted without changing the meaning of the respective CACs, the lexicalization and morphologization of such original periphrastic expressions is relatively low.

DOI 10.15460/auue.2025.98.1.347

4 Discussion

The data presented in the previous section demonstrate that CACs in Mokpe and Oroko match the features associated with the typologically-driven prototype of CACs – they do so both closely and with an equal regularity. Therefore, the CAC categories in the two languages may be viewed as similarly canonical. With regard to function, in an almost exception-less manner, Mokpe/Oroko CACs express motion-related actions and have domestic species as their addressees. They also tend to be monosemous rather than polysemous and are largely determined by the (very similar) ecosystems in which the Mokpe and Oroko communities live. With regard to form, primary CACs tend to be monosyllabic (or consist of identical monosyllabic segments replicated in series); exhibit a more consonantal than vocalic nature; contain extra-systematic sounds (both IPA and non-IPA) and sound combinations, and tolerate extra-systematic degrees of vocalic and consonantal length. In addition, primary CACs are monomorphemic (with replication constituting an expressive/phonetic rather than derivative/morphological device) and thus fail to exhibit inflections and derivations or draw on compounding. As all these formal properties are often violated by secondary CACs (which is also expected given typological data available; Andrason & Karani 2021), the CAC categories in Mokpe and Oroko are structurally opaque.

The evidence provided also confirms the phylogenetic hypothesis put forward on the basis of other languages and demonstrates the low cognancy of CACs in Mokpe and Oroko. The similarity of secondary CACs in both languages is low (especially for Oroko) and only a few cognate forms exist. Contrary to secondary CACs – and apparently violating one of the proposals concerning the cognancy of CACs – primary CACs exhibit a greater extent of similarity. However, this convergence, most likely, does not reflect shared ancestry – the identical CACs thus do not descend from earlier, for instance, proto-Sawabantu or proto-Bafawic-Bakweric forms. Rather, the resemblance attested is due to parallel developments, certainly similar yet independent. That is, by following common crosslinguistic strategies, which themselves are cognitively and/or iconically motivated, Mokpe and Oroko have separately developed CACs that look alike – a conclusion that stands in full agreement with the proposals concerning the phylogenetics of CACs (see section 2). Indeed, as is true of Ewe and Akan (Andrason

DOI 10.15460/auue.2025.98.1.347

& Gafatsi 2025), for all the cases of resemblance between Mokpe and Oroko CACs, equally similar forms exist in many other unrelated and geographically remote languages. Accordingly, despite the greater similarity of primary CACs than secondary CACs observed in Mokpe and Oroko, the hypothesis of the low cognancy of primary CACs (Daković 2006, Duah, Andrason & Antwi 2023) is, in our view, corroborated.

While our findings concur with the previous works dedicated to CACs and provide further support for the thesis of their low cognancy, a more general question emerges: Why are CACs in Mokpe, Oroko, and other languages so resistant to be transmitted along a language lineage? While more research is needed to offer a definitive solution to this issue, we think that the principal reason may concern the (preferred) non-arbitrariness and thus form-function transparency of CACs, especially the primary ones. That is, to preserve a direct relationship between their form and function – whether it is imitative/onomatopoeic (as is the case of several summonses) or motivated by other reasons (as is the case of dispersals) – and, inversely, to prevent CACs from acquiring more arbitrary profiles (something that typically occurs during the diachronic development of most language structures), CACs would regularly be "updated" and re-aligned to what speakers perceive as transparent, iconic, and motivated. While such "updating" is not necessary for many other lexical classes of sentence grammar (e.g. verbs, nouns, pronouns, numerals, adverbs, adjectives, conjunctions, and adpositions), it is critical for CACs because they are directed to animals. Indeed, our communication with these non-human species needs to be transparent because our "'interlocutor' is non-human and thus 'not in command of the [human] linguistic system'" (Andrason & Karani 2021: 4 citing Isačenko 1964: 95). (This however does not mean that CACs are never inherited across a language lineage or are always immune to be governed by diachronic laws operating in a language or language group. Both phenomena (may) occur although, we argue, they are less common than in sentence grammar.)

The same low cognancy across (closely) related languages and the same principle underlying the reluctance to be shared apply to onomatopoeias which, similar to CACs, are motivated and iconic. The comparison of the onomatopoeic categories in two languages of the Jen cluster (Adamawa), i.e. Dza and Mingang Doso, reveals that, at

DOI 10.15460/auue.2025.98.1.347

least in these varieties, onomatopoeias tend to be "reinvented" in the two branching varieties (Andrason & Benson 2023). Crucially, "this greater mutability of onomatopoeias […] may stem from their inherent expressiveness […] and iconicity" (Andrason & Benson 2023: 21). In other words, "the relatively direct relationship onomatopoeias entertain with extra-linguistic reality and the apparent need to imitate it, seems to motivate speakers to re-invent onomatopoeias over the course of the history of a language [family] to constantly match real-world sounds (as perceived by humans) with their linguistic representations" (Andrason & Benson 2023).

Additionally, the results of our research, together with data from other languages, for example Dogon (Andrason & Sagara 2024), suggest certain hierarchical correlations between the three formal types of CACs (primary, secondary, and borrowed) and the motion-related semantic types (summonses, dispersals, and directionals). Primary CACs tend to be used as summonses more commonly than dispersals, and only residually as directionals. In contrast, secondary and borrowed CACs tend to be used as directionals, less so as dispersals, and rarely as summonses. Inversely, directionals tend to be expressed with secondary and borrowed CACs; dispersals with the three formal types relatively equally; and summonses mainly with primary CACs. This can be captured by the following table:

Table 2. Correlations between the formal and semantic types of CACs

	Summonses	**Dispersals**	**Directionals**
Primary	Common	Semi-common	Uncommon
Secondary	Uncommon	Semi-common	Common
Borrowed	Uncommon	Semi-common	Common

5 Conclusion

The present article demonstrates that the categories of conative animal calls in Mokpe and Oroko comply with synchronic and diachronic tendencies typifying these types of constructions in the languages of the world. Primary CACs closely match the prototype of CACs and can be regarded as its canonical instantiations; secondary CACs may violate the prototypical features, which is also regular from a typological perspective. Furthermore, the comparison of Mokpe and Oroko CACs reveals the low cognancy of CACs in these two Sawabantu languages.

DOI 10.15460/auue.2025.98.1.347

The greater similarity exhibited by primary tokens than is the case of secondary ones – which is somewhat unexpected in a crosslinguistic context – does not stem from shared ancestry but is due most likely to parallel developments and the exploitation of fully motivated, typologically common strategies.

We propose that the main reason behind the reluctance of CAC constructions to be transmitted within a linguistic lineage is the principle of non-arbitrariness and form-function transparency governing the grammatical life of CACs, especially the primary ones. While this explanation coincides with what has been suggested to account for the similar low cognancy of onomatopoeias, more research on CACs in related languages is needed. We particularly need systematic diachronic studies dedicated to languages with well-established phonetic and morphological evolutionary rules (and a literary tradition extending over many centuries), such as Romance, Germanic, and Slavonic languages.

References

Aikhenvald, Alexandra Y. 2010. *Imperatives and commands*. Oxford: Oxford University Press.

Amha, Azeb. 2013. Directives to humans and to domestic animals: The imperative and some interjections in Zargulla. In Marie-Claude Simeone-Senelle and Martine Vanhove (eds.), *Proceedings of the 5th International Conference on Cushitic and Omotic languages*, Paris, 16–18 April 2008. Köln: Rüdiger Köppe Verlag. 211–229.

Andrason, Alexander. 2021. Polish borrowings in Wymysorys: A formal linguistic analysis of Germano-Slavonic language contact in Wilamowice. Reykjavík: University of Iceland PhD Dissertation. https://opinvisindi.is/handle/20.500.11815/2695

Andrason, Alexander. 2022. Conative animal calls in Xhosa: Testing the prototype. *Africana Linguistica* 28. 25–54. https://doi.org/10.2143/AL.28.0.3291207

Andrason, Alexander. 2023. The non-arbitrariness of some conative calls used to chase animals. *Linguistica Silesiana* 44. 73–103. https://doi.org/10.24425/linsi.2023.144824

Andrason, Alexander. 2024. Conative “kisses” in human-to-animal communication. *Lodz Papers in Pragmatics* 20(1). 103–129. https://doi.org/40.10.1515/lpp-2024-2001

DOI 10.15460/auue.2025.98.1.347

Andrason, Alexander & Pius W. Akumbu. 2024. Conative animal calls in Babanki and Bum (Grassfields languages of Cameroon). *KERVAN – International Journal of African and Asiatic Studies* 28(1). 3–28. https://ojs.unito.it/index.php/kervan/article/view/9525/8339

Andrason, Alexander, Anna Louisa Daigneault, Maya Hendrix, Rebecca Pizzitola & Anusha Somu. 2024. *A living database of conative animal calls*. Salem, Oregon: Living Tongues Institute for Endangered Languages. SSRN: https://ssrn.com/abstract=4899510. https://doi.org/10.5281/zenodo.12724966

Andrason, Alexander & Gabriel Kwasi Gafatsi. 2025. Typological, phylogenetic, and areal properties of conative animal calls in Ewe. *Yearbook of Oriental Studies* 78 (1). 5–37. https://doi.org/10.24425/ro.2025.154650

Andrason, Alexander & Michael Karani. 2021. Conative calls to animals: From Arusa Maasai to a cross-linguistic prototype. *Lodz Papers in Pragmatics* 37(1–2). 3–41. https://doi.org/10.1515/lpp-2021-0002

Andrason, Alexander, Onsho Mulugeta & Shimelis Mazengia. 2024. Conative animal calls in Macha Oromo: Function and form. *Linguistic Vanguard* 10. 199–207. https://doi.org/10.1515/lingvan-2023-0056

Andrason, Alexander & Admire Phiri. 2023. Talking to animals in a moribund language: Pragma-semantics, phonetics, and morphology of conative animal calls in Tjwao. *Linguistic Variation* 23(2). 1–25. https://doi.org/10.1075/lv.22008.and

Andrason, Alexander & Idrissa Sagara. 2024. The grammar of conative animal calls: The case of Togo-Teŋu Kan of Dourou (Dogon). *Asian and African Studies* 33(1). 48–96. https://doi.org/10.31577/aassav.2024.33.1.03

Atindogbé, Gratien G. 2012. On the typology of directional verbs in Bantu A (Barombi, Isubu, Mokpe, and Oroko). In Angelika Mietzner & Ulrike Claudi (eds.), *Directionality in grammar and discourse: Evidence from African Languages*. Rüdiger Köppe Verlag. 55–76.

Atindogbé, Gratien G. 2013. *A grammatical sketch of Mòkpè (Bakweri), Bantu A20*. African Study Monographs, Supplementary issue 45. 1–163. http://dx.doi.org/10.14989/171629.

Atta, Samuel Ebongkome. 1993. The phonology of Lukundu (Bakundu). Yaoundé: University of Yaoundé MA thesis.

Bate, Beatrice Bessem. 1987. The noun class system of Bakweri. Yaoundé: University of Yaoundé MA thesis.

DOI 10.15460/auue.2025.98.1.347

Brown, Cecil, Eric Holman, Søren Wichmann & Viveka Velupillai. 2008. Automated classification of the world's languages: A description of the method and preliminary results. *STUF – Language Typology and Universals* 61. 285–308.

Bynon, James. 1976. Domestic animal calling in a Berber tribe. In William McCormack & Stephen Wurm (eds.), *Language and man: Anthropological issues*. The Hague: Mouton. 39–66.

Che, Shu A. 1993. The noun class system of Oroko East. Yaoundé: University of Yaoundé MA thesis.

Connell, Bruce (ed.). 1997. *Mokpe (Bakweri)–English dictionary*. Köln: Rüdiger Köppe Verlag.

Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.

Croft William & Alan Cruse. 2004. *Cognitive linguistics.* Cambridge: Cambridge University Press.

Daković, Sybilla. 2006. *Interiekcje w języku polskim, serbskim, chorwackim i rosyjskim. Opis i konfrontacja.* Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Duah, Reginald Akuoko, Alexander Andrason & Mike Ofori Antwi. 2023. Phylogenetics of conative animal calls: Asante, Bono, and Fante. "Phylogenetics of conative animal calls: Asante, Bono, and Fante." Talk at the Semantics Colloquium, Institute für Linguistik, Goethe Universität Frankfurt, 5 May 2023.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2023. *Ethnologue: Languages of the world. Twenty-sixth edition*. Dallas: SIL International. Online version: http://www.ethnologue.com.

Ebobissé, Carl. 1989. Dialectrometrie lexical del parlers Sawabantu. *Journal of West African Languages* 19(2). 57–66.

Ebobissé, Carl. 2015. *Sawabantu: Eine vergleichende Untersuchung der Küstensprachen Kameruns (Bantu A.20 und A.30)*. Frankfurt: Peter Lang.

Eide, Phyllis J. 2008. Recruiting participants. In Lisa Given (ed.), *The Sage encyclopedia of qualitative research methods*. Los Angeles: Sage. 743–745.

Eren, Hasan. 1952. Zurufe an Tiere bei den Türken. *Ural Altaische Jahrbücher* 23. 134–137.

Evans, Vyvyan & Melanie Green. 2006. *Cognitive linguistics: An introduction*. Edinburgh: Edinburgh University Press.

Friesen Dan. 2002. Oroko orthography development: Linguistic and sociolinguistic factors. Grand Forks: University of North Dakota MA thesis.

DOI 10.15460/auue.2025.98.1.347

Friesen, Dan & Lisa Friesen. 2001. Extendibility survey of Oroko. Ms.

Friesen, Lisa. 2002. Valence change and Oroko verb morphology (Mbonge dialect). Grand Forks: University of North Dakota MA thesis.

Gardani, Francesco. 2020. Morphology and contact-induced language change. In Anthony Grant (ed.), *The Oxford handbook of language contact*. 96–122. Oxford: Oxford University Press

Gensler, Orin David. 1980. Verbal morphotonemics of Bakweri. Berkeley: University of California MA thesis.

Gensler, Orin David. 1981a. The phonological status of downstep in Bakweri. In Danny K. Alford, Karen Ann Hunold, Monica A. Macaulay, Jenny Walter, Claudia Brugman, Paula Chertok, Inese Čivkulis & Marta Tobey (eds.), *Proceedings of the Seventh Annual Meeting of the Berkeley Linguistics Society*. 50–67. https://doi.org/10.3765/bls.v7i0.2088

Gensler, Orin David. 1981b. The representation of vocalic object-affixes and tense-affixes in the Bakweri verb. *Studies in African Linguistics, Suppl.* 8. 38–43.

Germanovič, Aleksandr J. 1954. Слова клича и отгона животных в русском языке. *Известия Крымского Педагогического Института* 19. 306–320.

Guthrie, Malcolm. 1967/71. *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages, 4 vols*. Letchworth & Brookfield: Gregg International.

Hammarström, Harald. 2019. An inventory of Bantu languages. In Mark Van de Velde, Koen Bostoen, Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*. New York: Routledge. 17–78. https://doi.org/10.1017/s0022226714000127

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank. 2023. *Glottolog 4.8*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://glottolog.org

Hawkinson, Ann. 1986. Bakweri verb morphology. Berkeley: University of California PhD dissertation.

Heine, Bernd. 2023. *The grammar of interactives*. Oxford: Oxford University Press.

Hombert, Jean-Marie. 1973. Speaking backwards in Bakwiri. *Studies in African Linguistics* 4. 227–236.

Isačenko, Alexander. 1964. On the conative function of language. In Josef Vachek (ed.), *A Prague School reader in linguistics*. Bloomington: Indiana University Press. 88–97.

DOI 10.15460/auue.2025.98.1.347

Janda, Laura. 2015. Cognitive linguistics in the year 2015. *Cognitive Semantics* 1. 131–154. https://doi.org/10.1163/23526416-00101005

Jarring, Gotthard. 1941. Zurufe an Tiere bei den Osttürken. In Gotthard Jäschke (ed.), *Festschrift Friedrich Giese aus Anlass des siebzigsten Geburtstags*. 46–52. Berlin: Deutsche Gesellschaft für Islamkunde.

Jääskeläinen, Anni. 2021. Eläinten omat sanat: Eläimille suunnatut interjektiot ihmisen ja eläimen vuorovaikutuksen osina. In Rea Peltola, Anni Jääskeläinen & Katariina Harjunpää (eds.), *Kieli ja elain. Vuorovaikutusta ja kielioppia monilajisissa yhteisöissä*. Helsinki: Suomalaisen Kirjallisuuden Seura. 133–184.

Kagaya, Ryohei. 1992a. *A classified vocabulary of the Bakueri language*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa (ILCAA).

Kagaya, Ryohei. 1992b. A tonal system of nouns, possessives and adjectives of the Bakweri language—on two types of basic tones. *Bantu Linguistics (ILCAA), Studies in Cameroonian and Zairean Languages* 3. 231–268.

Kagaya, Ryohei. 1992c. A tonal system of verbs of the Bakweri language. *Bantu Linguistics (ILCAA), Studies in Cameroonian and Zairean Languages* 3. 95–230.

Mabian, Mercy. 2010. *Oroko textbook drive project: A continual project*. https://orokousa.org/wp-content/uploads/2015/03/Books_for_Knowledge.pdf

Marlo, Michael & David Odden 2007. The exponence of TAM in Bakweri. *SOAS Working Papers in Linguistics* 15. 19–31.

Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

Mbongue, Joseph. 2000. A rapid appraisal survey of Oroko. Yaoundé: CABTAL.

Meakins, Felicity. 2011. *Case-marking in contact: The development and function of case morphology in Gurindji Kriol*. Amsterdam: John Benjamins.

Mokake, Levi E. 2016. The lexical phonology of Mokpe. Buea: University of Buea MA thesis.

Mutia, Babila. 2005. Performer, audience, and performance context of Bakweri pregnancy rituals and incantations. *Cahiers d'Études Africaines* 45(1). 217–237. https://doi.org/10.4000/etudesafricaines.4942

Neh, Monikang Evelyn. 1989. Phonology Of Mokpwe. Yaoundé: University of Yaoundé MA thesis.

DOI 10.15460/auue.2025.98.1.347

Ngoisah, Lyonga A. 2002. The morphology of the noun phrase and the verb phrase in Mokpe in relation to orthography. Buea: University of Buea MA thesis.

Nurse, Derek & Gérard Philippson. 2003. Towards a historical classification of the Bantu languages. In Derek Nurse & Gérard Philippson (eds.), *The Bantu languages*. London: Routledge. 164–181. https://doi.org/10.1017/s0041977x05490278

Pakendorf, Brigitte. 2009. Intensive contact and the copying of paradigms: An Ėven dialect in contact with Sakha (Yakut). *Journal of Language Contact* 2(2). 85–110.

Poyatos, Fernando. 1993. *Paralanguage: A linguistic and interdisciplinary approach to interactive speech and sounds*. Amsterdam: John Benjamins.

Poyatos, Fernando. 2002. *Nonverbal communication across disciplines. Volume II: Paralanguage, kinesics, silence, personal and environmental interaction.* Amsterdam: John Benjamins.

Saumure, Kristie & Lisa M. Given. 2008. Non-probability sampling. In Lisa Given (ed.), *The Sage encyclopedia of qualitative research methods.* Los Angeles: Sage. 562–563.

Schötz, Susanne. 2020. Phonetic variation in cat-human communication. In Ramiro Pastorinho & Ana Catarina Sousa (eds.), *Pets as sentinels, forecasters and promoters of human health.* Cham: Springer Nature. 319–347.

Schulthess, Friedrich. 1912. *Zurufe an Tiere im Arabischen*. Berlin: Abhandlungen der Akademie der Wissenschaften.

Siatkowska, Ewa. 1976. *Zachodniosłowiańskie zawołania na zwierzęta: stan obecny, funkcje historyczne, stosunek do systemu językowego*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Tanda, Vincent A. & Ayu'nwi N. Neba. 2005. Negation in Mòkpè and two related Coastal Bantu languages of Cameroon. *African Study Monographs* 26(4). 201–219. https://doi.org/10.14989/68243

Treis, Yvonne. 2023. Interjections in Kambaata. Paper presented at the Afrikanist*innentag 2023, Leipzig, 4–6 May 2023. https://hal.science/hal-04092140

Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: John Benjamins.

Wichmann, Søren 2007. Guidelines for preparing 40-word lists for languages to be included in the ASJP database. https://asjp.clld.org/static/Guidelines.pdf

DOI 10.15460/auue.2025.98.1.347

Wichmann, Søren, André Müller, Viveka Velupillai, Annkathrin Wett, Cecil Brown, Zarina Molochieva, et al. 2012. The automated similarity judgment program (ASJP): The ASJP database (version 15). http://email.eva.mpg.de/~wichmann/languages.htm. https://hdl.handle.net/11858/00-001M-0000-000E-8029-8

Wichmann, Søren, Eric Holman & Cecil Brown (eds.). 2022. *The ASJP database (version 20)*. https://asjp.clld.org

Wierzbicka, Anna. 2003. *Cross-cultural pragmatics: The semantics of human interaction*. Berlin: De Gruyter.

DOI 10.15460/auue.2025.98.1.370

Number marking on nouns and adjectives of Sidaama

Anbessa Teferra

Tel Aviv University
teferra@tauex.tau.ac.il

Abstract
This paper aims to systematically analyze the formal marking of numbers on common nouns and adjectives of Sidaama, a Highland East Cushitic (HEC hereafter) language spoken in Ethiopia. There are three formal categories of number in Sidaama: basic form, singulative, and plurative. The basic form is a form of a noun that is not marked for number; a singulative denotes a single referent, and a plurative marks multiple referents. The singulative of nouns is usually marked either by *-ʧ-o* ~ *-ʧʧ-o*. There are also a few nouns that contain the formatives *-ʃ-o*, *-ʧ-o*, and *-k-o* as markers of the singulative. Such forms arose from the merging of the final consonant of the nominal root with the initial *-ʧ* of the singulative. Unlike nouns, singulative is marked only on a handful of adjectives. Five types of nominals and six types of adjectival pluratives are identified. All of them are suffixal, and two of them involve the copying of a root-final consonant. Sidaama has inherently singular nouns that usually denote humans and higher animals. In such nouns, their basic form marks a single referent. The language also contains inherently plural nouns that refer to a collective of entities or a group. Few nouns are transnumeral; in their basic form, they can refer to either a single entity or plural entities.

Keywords: Sidaama, Cushitic, number marking, basic form, singulative, plurative

1 Introduction

This section provides general information about the Sidaama language and discusses the study's objectives, significance, dataset, and research structure.

DOI 10.15460/auue.2025.98.1.370

1.1 Background information

Sidaama is one of the HEC languages comprising Hadiyya, Kambaata, Alaaba, K'abeena, Gedeo, and Burji. The autoglottonym of Sidaama is *Sidaam-u ʔafoo* [lit. "Sidaama-of mouth"] or *Sidaam-u k'aale* [lit. "Sidaama-of word"] while its ethnonym is Sidaama. Although native speakers label their language *Sidaamu ʔafoo*, the name adopted in this paper and elsewhere is Sidaama since it has a broader use and currency in the linguistic literature. According to the 2007 national census of Ethiopia, the number of Sidaama mother-tongue speakers was 2,925,171 (Central Statistical Authority 2010: 200). The Sidaama language is spoken in southern in Ethiopia within the Administrative Region of Sidaama. Small enclaves of Sidaama populations are found in south Ethiopia among the neighboring Arsi Oromo, Guji Oromo, and Gedeo. Other ethnic groups, such as Amhara, Oromo, Wolaitta, Kamabaata, the Gurage, Silte, etc., live among the Sidaama, mainly in towns.[1] The capital city, Hawaasa,[2] is a microcosm of various ethnic groups and is known as 'little Ethiopia'. Except for Omotic-speaking Wolaitta in the west, the Sidaama are bounded by speakers of Cushitic languages: Guji Oromo in the east, Arsi Oromo in the north, and Gede'uffa speakers in the south.

Since 1992, the language has exhibited tremendous growth. It is a language of instruction within the Administrative Region of Sidaama at the first cycle of primary level (Grade 1–4) in rural and urban schools. It is taught as a subject in all grades. In addition, it is used as a language of administration and jurisdiction and, to some extent, as a language of mass media in the Administrative Region of Sidaama. In 2014, a B.A. program in Sidaama linguistics and literature was launched at Hawassa University.

Typologically, Sidaama is a head-final language with canonical SOV word order. Hence, embedded clauses occur before main clauses while modifiers precede their heads. Thus, adjectives, genitives, and relatives precede a nominal head, while adverbs precede their verbal heads. Nouns and their modifiers agree in case and gender, while

1 Gurage is a cover term that refers to various ethnic groups and language varieties mainly found in the Gurage Zone.

2 The correct name of the city in the Sidaama language is Hawaasa, and I have adopted this correct form. However, it is incorrectly written as Hawassa.

DOI 10.15460/auue.2025.98.1.370

number agreement is restricted. Sidaama is a suffixing language save for the prefix *di-* that marks negation.

1.2 Objectives and scope of the study

Sidaama is one of the HEC languages that is relatively well studied in its group.[3] Nevertheless, various grammatical elements necessitate a detailed analysis to get a more precise and accurate picture, and one of these is number marking on nouns and adjectives. There are various types of research on Sidaama that touch upon number marking. This includes Kawachi 2007, Anbessa 2012, and Yri 2011. However, some elements were not analyzed in detail, and in other cases, amendments are necessary. Hence, in this article, number-marking on nouns and adjectives will be analyzed in detail to get a clearer picture.

This research will be restricted to number marking on nouns and adjectives. Hence, it will not involve number marking for other word classes, such as verbs and pronouns.

1.3 Dataset and structure of the research

Since the current study is primarily lexical and morphological, the tools employed to get the necessary data are written corpus and introspection. In this article, the written corpus consists of materials on number marking in Sidaama, HEC, and Cushitic languages and it includes Anbessa (2012; 2014), Kawachi (2007), Treis (2008; 2014), and Kramer & Anbessa (2020). In addition, the two dictionaries on Sidaama, i.e., Gasparini (1983) and Shimelis (2007), were consulted. Introspection was used as a supportive tool in the study because the writer of this article is a native speaker of Sidaama. Nevertheless, in the future, an effort will be exerted to use spoken data and written corpus in various Sidaama books to broaden the research.

The article contains four sections. The first introductory section is followed by the second section, the article's main body. This includes the linguistic analysis of the basic form, singulative, and plurative in nouns and adjectives of Sidaama. In the third section, the functions of number marking will be discussed. In this section, inherently singular, inherently plural, and transnumeral nouns will be dealt with. In the fourth section, the findings will be summarized.

3 I think Kambaata also enjoys the same privilege thanks to extensive linguistic research conducted on the language by Yvonne Treis (i.e. 2008; 2014).

DOI 10.15460/auue.2025.98.1.370

The transcription adopted in this paper is phonemic and adheres to the standard IPA. However, wherever relevant, a phonetic transcription is provided. The transcription adopted here differs from the standard IPA only in representing long vowels and geminates, which are transcribed by doubling a vowel and a consonant.

2 Number in Sidaama

Sidaama nouns are marked for case, gender, and number. A single vocalic suffix marks the three prominent cases, comprising nominative, accusative, and genitive. There are also oblique cases that are usually formed on a genitive base. Nouns distinguish two grammatical genders: masculine and feminine. According to Yri (2012: 262), "prototypical nouns have inherent class membership in one of the two nouns (gender) classes: k-class (M) or t-class (F)." Except for some lexemes denoting kinship terms and higher animals, the gender of all other nouns is not predictable and, hence, is assigned arbitrarily. Thus, gender is mainly realized through agreement features on nouns, adjectives, and verbal endings.

Three formal categories of number are identified in Sidaama across most morphosyntactic contexts. These are basic form, singulative, and plurative.[4] Few nouns have all three categories of number, while other nouns may have only one (the basic form) or a combination of two number categories. Among closely related HEC languages, the number in Kambaata was discussed formally and in detail by Treis (2008; 2014). Hence the terminologies for formal categories of number (such as "basic form", "singulative", and "plurative"), in this paper were adopted from Treis, and the grammatical analysis of number in Sidaama has benefited a lot from her research. That being the case, number in Sidaama was analyzed based on language-internal data. In addition, it must be noted that although Kambaata and Sidaama are related HEC languages, they also have many differences.

2.1 Basic form

Treis (2008: 131) defines the basic form as "[…] the form of the noun which is formally unmarked for number". Thus, the basic form

4 Regarding number categories in Kambaata, Treis (2014: 2) writes, "In order not to confuse form and meaning, the terms 'basic form', 'singulative', and 'plurative' are used with respect to the form of a noun".

DOI 10.15460/auue.2025.98.1.370

in Sidaama has no number marking, and it consists of a noun stem that is followed by a terminal vowel[5] which marks the accusative. Yri (2012: 262) calls the basic form "neutral", while Kawachi (2007: 344) and Anbessa (2012: 22; 24) label it "unmarked". Anbessa (2014: 28) also calls it "collective", as does Gasparini (1983), and many such forms in his dictionary entries are abbreviated as "coll.". Examples of basic forms in Sidaama are nouns such as *badal-a* 'maize', *dog-o* 'road', *min-e* 'house', etc. In all the three words, there is no marking for number. For instance, in the word *badal-a* 'maize', the nominal root[6] is *badal-* and the terminal vowel is *-a*. All the three words are citation forms, and their terminal vowels mark the accusative.

There are few freely standing (independent) basic forms, as seen in the three examples above. Instead, the language uses nominal roots, reconstructed based on their singulative, plurative, or both (see Table 1 and Table 2). An asterisk marks such reconstructed roots.

Table 1. Roots

Nominal root	Singulative	Plurative
**haiss-*	*haiss-i-ʧ ʧ-o* 'a blade of grass'	*haiss-o* 'grass' (PL)
**kin-*	*kin-ʧ-o* 'stone'	*kin-n-a* 'stones'
**k'up'p'-*	*k'up'p'-i-ʧʧ-o* 'egg'	*k'up'p'-e* 'eggs'
**reek'k'-*	*reek'k'-ii-ʧʧ-o* 'obsidian'	*reek'k'-a* 'obsidians'

2.2 Singulative

A singulative denotes a single referent and can be marked both on nouns and adjectives. Nonetheless, marking a singulative on nouns is quite frequent, while in the case of adjectives, it is very restricted.

2.2.1 Singulatives of nouns

The singulative of nouns is marked mainly by *-ʧ-o* ~ *-ʧʧ-o* and, in a few cases, by other sets of allomorphs.

5 The label "terminal vowel" is borrowed from Azeb Amha (2017: 822) and is common in Omotic languages. Although Sidaama is not an Omotic language, the label "terminal vowel" suits my analysis well, which is why it is adopted in this paper.

6 In this article, the term 'root' refers to a bare morpheme that is devoid of grammatical affixes.

DOI 10.15460/auue.2025.98.1.370

(i) Singulatives marked with *-ʧ-o* ~ *-ʧʧ-o*

The singulative of nouns is marked mainly by *-ʧ-o* or *-ʧʧ-o*. The actual singulative formatives are *-ʧ* or *-ʧʧ*, while *-o* is a terminal vowel that marks an accusative and provides the citation form. For instance, in the word *kin-ʧ-o* 'stone', *-o* is the terminal vowel, and this word, being accusative, will appear in the object position. When the same word appears in the subject position, the terminal vowel will be replaced by *-u*, which marks nominative in Sidaama and becomes *kin-ʧ-u* 'stone' (nom). The singulative *-ʧ-o* is suffixed to nominal roots that end in sonorants. Sidaama has seven sonorants: *l*, *m*, *n*, *ɲ*, *r*, *w*, and *j*. However, nominal roots to which a singulative is suffixed end only in three of them: *n*, *r*, and *l*, while nominal roots with final *m*, *ɲ*, *w*, and *j* were not attested (see Table 2).

Table 2. Nouns marked by the singulative *-ʧ-o*

Nominal root	Singulative	Gloss
**fur-*	*fur-ʧ-o*	'a small opening'
**dar-*	*dar-ʧ-o*	'leaf'
**hoon-*	*hoon-ʧ-o*	'a juniper tree'
**k'amal-*	*k'amal-ʧ-o*	'a small monkey'
**k'iissan-*	*k'iissan-ʧ-o*	'spider'

The above roots are nominals, to which are attached singulatives. However, other roots display an interesting morphological feature: the same root can be shared by both nouns and verbs, and many roots in Sidaama yield noun-verb pairs. For instance, the same root can suffix a singulative and become a noun, while it can also attach an inflectional suffix and become a verb. In this article I label such kind of root "pre-categorial", as suggested by Ronny Meyer in personal communication. Derivations that are based on it are illustrated in tables (3) and (4). If singulative suffixation to a root yields a noun, it has a derivative function in addition to its number-marking major function.

DOI 10.15460/auue.2025.98.1.370

Table 3. Pre-categorial roots with nominal outputs

Pre-categorial root	Singulative	Gloss
**da'mul-*	*da'mul-ʧ-o*	'worm'
**darar-*	*darar-ʧ-o*	'flower'
**lal-*	*lal-ʧ-o*	'early/first fruit'

Table 4 shows how an inflectional formative derives verbs from the same roots listed in Table 3.

Table 4. Pre-categorial roots with verbal outputs

Pre-categorial root	Verbal form	Gloss
**da'mul-*	*da'mul-í*	'It became worm eaten.'
**darar-*	*darar-í*	'It blossomed.'
**lal-*	*lal-í*	'It grew.'

In Table 4, by attaching the 3M.SG of inflectional suffix *-í* to the same pre-categorial root *darar-*, one gets the verbal form *darar-í* 'it blossomed'. If both nouns and verbs share the same root, which is the basic? Regarding this dilemma in a related HEC language, Kambaata, Treis (2008: 162) writes:

> "At the present state of knowledge, one cannot decide which lexeme is to be considered as the basic one and which one is converted, i.e., 'zero-derived'. Determining whether a stem like *caakk-* is basically a noun, *caakk-á* 'light', or a verb, *caakk-ú* 'to give light (INF)', is not possible as both only differ with respect to their inflectional behaviour. One lexeme adds endings of the nominal declension M1, and the other one adds verbal endings."

In addition, after examining the relation between verbs and adjectives, Treis (2008: 269) writes, "[t]he word class of lexical stems in Kambaata is **not pre-determined**; one and the same stem may serve as a verb or adjective […]." (bolding by Treis).

I support the analysis of Treis because the same phenomenon is observed in Sidaama. Many pre-categorial roots can yield verbs and nouns, and deciding the basic word class of such roots is challenging. There are even more complex cases because, from the same pre-categorial root, it is possible to derive a verb, a noun, and an adjective. For instance, the root *goww-* yields an inchoative verb *goww-í* 'he became foolish' (by suffixing *-í* = 3M.SG.PERF), an adjective *goww-a*

DOI 10.15460/auue.2025.98.1.370

'foolish' (by attaching the adjectivizer *-a*[7]), and a noun *goww-imma*, 'foolishness' (via suffixation of the nominalizer *-imma*).

The second singulative allomorph, i.e., *-ʧʧ-o*, is suffixed to nominal roots that end in a single obstruent, a geminate, or a consonant cluster. In such forms, an epenthetic *-i* is inserted between the nominal root and the singulative formative to disallow impermissible consonant sequences (see Table 5).

Table 5. Nouns marked by the singulative *-ʧʧ-o*

Nominal root	Singulative	Gloss
**haant'-*	*haant'-ii-ʧʧ-o*[8]	'enset fibre'
**haiss-*	*hais-i-ʧʧ-o*	'a blade of grass'
**hak'k'-*	*hak'k'-i-ʧʧ-o*	'wood/tree'
**ʧ'e?-*	*ʧ'e?-i-ʧʧ-o*	'bird'
**wiliː?l-*	*wiliː?l-i-ʧʧ-o*	'lamb'
**woʃ-*	*woʃ-i-ʧʧ-o*	'dog'

(ii) Singulatives that contain other formatives

Some nouns contain formatives such as *-ʃ*, *-ʧ*, and *-k* as markers of the singulative synchronically. Such singulatives arose because of the historical morphophonological processes of Sidaama, as will be explained immediately below.

(a) Singulatives marked with *-ʃ-o*

There are a handful of singulatives that contain *-ʃ* synchronically. It appears that first the *-ʧ* of the singulative became *-ʃ* through some historical assimilatory process. Then the root-final consonants completely assimilated to the *-ʃ* of the singulative resulting in *-ʃʃ*. Thus, in such forms, the actual singulative marker is *-ʃ* and not *-ʃʃ*. This is because the first *-ʃ* is a root-final consonant which is not part of the singulative marker. This assimilatory process was observed first by Moreno (1940: 80). The evidence for this can be gleaned from basic forms whereby the root-final consonant can be observed. Additionally, in three of the words, the final consonant of the nominal root in plurative forms provides evidence (Table 6).

7 The formative *-a* also marks an infinitive in Sidaama. Thus, *goww-a* can also be translated as 'to become foolish'.

8 In some nouns, the epenthetic vowel is lengthened, and there is no explanation for the time being.

DOI 10.15460/auue.2025.98.1.370

Table 6. Singulatives marked with *-ʃ-o*

Basic form	Singulative	Plurative
ʔagad-a	*ʔagaʃ-ʃ-o* 'stem of maize'	----------
farad-o	*faraʃ-ʃ-o* 'horse'	*farad-da* 'horses'
galad-o	*galaʃ-ʃ-o* 'monkey'	*galad-da* 'monkeys'
hamas-o	*hamaʃ-ʃ-o* 'snake'	----------
wees-e	*weeʃ-ʃ-o* 'enset'[9]	----------
**rumud-*	*rumuʃ-ʃ-o* 'root'	*rumud-d-a* 'roots'

In the above examples, the root-final *d* or *s* has merged with the singulative *ʧ*, resulting in **ʃʃ*.[10] As shown below in examples 1) and 2), many of the above basic forms are viewed as collective or a single group, which is observed in their singular verbal inflection.

(1) *bero* *farad-u* *buraak'-í*
Yesterday horses-NOM gallop-3M.SG.PERF
'Yesterday the horses galloped.'

(2) *galad-u* *hand-o* *ʃorr-í*
monkeys-NOM ox-NOM chase-3M.SG.PERF
'The monkeys chased the ox.'

Unlike the singulative and plurative, the basic form lacks number-marking formatives. Nevertheless, semantically, some basic forms can convey a plural meaning although they are not plurative.

(b) Singulatives marked with *-ʧ-o*

The singulative *-ʧʧ* that occurs in nouns in Table 5, and additional nouns such as *k'ubb-i-ʧʧ-o* 'finger', *mik'-i-ʧʧ-o* 'bone', *woʃ-i-ʧʧ-o* 'dog', etc., is an inherent and indivisible one. There are some nouns (see Table 7 below) that apparently contain *-ʧ-ʧ* as a singulative marker. However, in these nouns too, singulative is marked only by *-ʧ*. The first *ʧ* is actually a root-final consonant that has assimilated to the *-ʧ* of the singulative. In such nouns, the synchronic *-ʧ-ʧ* results from a diachronic morphophonological process whereby the root-final con-

9 The word 'enset' (Ensete ventricosum) refers to a plant that, among other names, is also known as 'false banana'.

10 Moreno (1940: 80), Anbessa (2000: 35; 2014: 27), and Kawachi (2007: 347) discussed this and similar morphophonological processes in Sidaama. A similar process is also observed in Kambaata and was explained by Treis (2008: 133).

DOI 10.15460/auue.2025.98.1.370

sonant (mostly *b* and, in a few cases, *d*) completely assimilated to *-ʧ* of the singulative resulting in *-ʧ-ʧ*. Again, this evidence is gleaned from the basic forms and, in two cases, from pluratives whereby the diachronic root-final consonant surfaces.

Table 7. Singulatives marked with *-ʧ-o*

Basic form	Singulative	Plurative
ʔawaad-o	*ʔawaʧ-ʧ-o* 'tanner'	----------
ʔibiib-e [*ʔiwiiw-e*]	*ʔibiiʧ-ʧ-o* 'louse'	----------
dagub-a [*daguw-a*]	*dagutʃ-ʧ-o* 'cedar tree'	----------
gereeb-o [*gereew-o*]	*geretʃ-ʧ-o* 'sheep'	*gereeb-ba* 'sheep'
gulub-e [*guluw-e*]	*gulutʃ-ʧ-o* 'knee'	*gulub-b-a* 'knees'

In Table 7, a pervasive weakening rule applies in four basic forms whereby an intervocalic */b/* becomes [*w*]. In the above examples, although *ʔawaad-o* and *gereeb-o* are basic forms, they convey a plurative meaning of 'tanners' and 'sheep' (PL), respectively. Thus, they are plurative in meaning but not in form. However, the basic form *gereeb-o* 'sheep' has a full-fledged plurative *gereeb-ba* 'sheep' (PL).

(c) Singulatives marked with *-k-o*

A handful of Sidaama singulatives contain *-k-o*. In such nouns, *-ʧ* of the singulative assimilated completely to the root-final *k* of the basic form, resulting in a geminate, i.e., *-k-k*. Again, in such forms the actual singulative marker is *-k* and not *-kk*. This is because the first *-k* is a root-final consonant which is not part of the singulative marker. It is the second *-k* that marks singulative and which surfaced after *-ʧ* of the singulative assimilated to the root-final *k*. In such nouns, the basic forms provide clear evidence whereby the diachronic root-final consonant *k* surfaces. In Table 8 I have provided phonetic forms because a lenition rule applies whereby intervocalic */k/* becomes [*h*].

Table 8. Singulatives marked with *-k-o*

Basic form	Singulative
ʔodak-e [*ʔodahe*]	*ʔodak-k-o* 'sycamore'
beeddak-e [*beeddah-e*]	*beedak-k-o* 'star'
futak-e [*futah-e*]	*futak-k-o* 'mole'
heyyak-e [*heyyahe*]	*heyyak-k-o* 'partridge'
yemak-e [*yemah-e*]	*yemak-k-o* 'rat'

DOI 10.15460/auue.2025.98.1.370

Sidaama contains the noun *muduk-k-o* 'navel' that synchronically contains *-k* as a singulative formative. Based on its singulative, I will propose **muduk-* as the nominal root of the basic form. An almost identical basic form is attested in two related HEC languages. Treis (2008: 133) lists **mudug-* as a root[11] in Kambaata based on *mudug-i-chch-ú*, while Hudson (1989: 104) lists the singulative *mudug-iccu* for Kambaata. According to Schneider-Blum (2007: 446) in Allaba, the word for 'navel' is *mudug-iccú*.

Kawachi (2007: 347) considers forms such as *-ečč̌o*, *-aššo*, *-uššo*,[12] and *-akko* as additional allomorphs of the singulative and lists the following singulative words: *ger-ečč̌o* 'sheep', *far-aššo* 'horse', *rum-uššo* 'root', and *fut-akko* 'mole', etc. among others. According to his analysis, only the first three segments such as *ger-*, *far-*, *rum-*, and *fut-* form the nominal root, while the remaining elements, such as *-ečč̌o*, *-aššo*, *-uššo*, and *-akko* are singulatives.

According to my analysis, the singulative allomorphs are *-ʧ*, *-ʃ*, and *-k* and not geminates such as *-ʧʧ*, *-ʃʃ* and *-kk*. This is because the first segments are actually root-final consonants which do not take part in the marking of singulative. Moreover, I do not support the position of Kawachi's morphemic cuts. My analysis differs from that of Kawachi because I exclude the initial vowels *-e*, *-a*, *-u*, and the segments *č*, *š*, and *k*, which are adjacent to the vowels he listed as part of the singulative allomorph. According to my analysis, these initial vowels and the adjacent consonants are part of the nominal root and must not be included in the singulative allomorphs. For instance, in synchronic nouns such as *gereʧʧo* 'sheep', *faraʃʃo* 'horse', *rumuʃʃo* 'root', and *futakko* 'mole' their presumed diachronic forms are **gereb-ʧo*, **farad-ʧo*, **rumud-ʧo* and **futak-ʧo*. Thus, segments that Kawachi lists as part of the singulative allomorphs are, in fact, the last vocalic and consonantal elements of the nominal root, as they can be seen in the reconstructed forms.

2.2.2 Singulatives of Adjectives

Sidaama has very few "genuine" (underived) adjectives such as *danʧa* 'good', *faajja* 'pretty', *ʃiima* 'little', etc. Instead, derivational processes enable the creation of numerous adjectives from verbs and

11 Treis uses the label "basic form".

12 The IPA equivalent of *č* is *ʧ* and that of *š* is *ʃ*.

DOI 10.15460/auue.2025.98.1.370

nouns. Sidaama adjectives share grammatical features such as case, gender, and number with nouns. Nevertheless, there are also differences. Regarding number, a noun can have three forms (basic form, singulative, and plurative) while adjectives maximally have two (basic form and plural or singulative and plural). As far as gender is concerned, nouns have an inherent gender (masculine or feminine), while adjectives do not. Gender-marked adjectives always occur in pairs as in: *matʃtʃaraar-aam-o* 'mad/crazy' (M) and *matʃtʃaraar-aam-e* 'mad/crazy' (F).

Adjectives can have an attributive or a predicative function. An attributive adjective precedes a noun that it modifies and agrees with it in case, gender, and number, as shown in examples 3a) and 3b) below.

(3a) *busul-u* *man-tʃ-i* *wot'e*
clever-NOM.M man-SGV-NOM.M money
di-fintʃ'-anno
NEG-scatter-3M.IMPERF
'A clever person will not scatter (waste) money.'

(3b) *busul-laadd-u* *man-oot-i* *wot'e*
clever-PL-NOM.M man-PL-NOM.M money
di-fintʃ'-i-tanno
NEG-scatter-3PL.IMPERF
'Clever men will not scatter (waste) money.'

Both adjectives and nouns can be used predicatively and take an identical copula: *ho* (M) and *-te* (F) as shown in ex. 4) and 5).

(4) *beett-u* *busule-ho*
boy-NOM.M CLEVER-COP.M
'The boy is clever.'

(5) *kun-i* *hando-ho*
M.PROX-NOM.M OX-COP.M
'This is an ox.'

Although adjectives and nouns take the same copula, they also differ. For instance, when modifiers precede predicative adjectives, they

DOI 10.15460/auue.2025.98.1.370

take either the copula *ho* (M) or *-te* (F), while predicative nouns take the copula *-ti*.

There are differences between adjectives and nouns regarding singulative: singulative is extensively marked on nouns and hardly on adjectives. Singulative is marked by *-ʧo* ~ *-ʧʧo* only in a few adjectives, as shown in Table 9.

Table 9: The singulatives of adjectives marked by *-ʧo* ~ *-ʧʧo*

Root	Singulative of adjectives
**but'-* 'become poor'	*but'-i-ʧʧ-o* 'poor'
gura 'left'	*gura-ʧʧ-o* 'left-handed'
**hara?m-*[13] 'be short'	*haran-ʧ-o* 'short'

In Table 9, the adjectives were derived by suffixing *-ʧʧo* and *-ʧo* to roots. As was mentioned earlier, here, too, it is impossible to determine the basic word class of a root such as *but'-* 'become poor.' This is because attaching the appropriate derivational suffixes makes it possible to derive different word classes. Thus, from *but'-* 'become poor', a singulative adjective can be derived by suffixing *-ʧʧ-o* (*but'-i-ʧʧ-o* 'poor') or the plurative adjective by attaching *-ane* (*but'-ane* 'poor ones'), a noun by suffixing *-ima* (*but'-ima* 'poverty'), and an inchoative verb via the 3M.SG.PERF suffix *-í* (*but'-í* 'he became poor').

Similarly to nouns, few adjectival singulatives contain *-ʃ* synchronically. It seems that in adjectives too, first the *-ʧ* of the singulative became *-ʃ* through some historical assimilatory process. Then *d*, i.e., the final consonant of the adjectival formative *-ad* have completely assimilated to the *-ʃ* of the singulative resulting in *-ʃ-ʃ*.

Table 10: The singulatives of adjectives marked by *-ʃ-o*

Pre-categorial Root	Adjectival stem	Singulative adjective
**beeb-* 'be closed'	*beeb-ad-*	*beeb-aʃ-ʃ-o* 'closed'
**ʤaaw-* 'become thin'	*ʤaaw-ad-*	*ʤaaw-aʃ-ʃ-o* 'thin'
**k'iid-* 'become cold'	*k'iid-ad-*	*k'iid-aʃ-ʃ-o* 'cold, dull'
**ʃaal-* 'be slim'	*ʃaal-ad-*	*ʃaal-aʃ-ʃ-o* 'slim'

13 Although 'to be short' is synchronically based on *hara?m-*, its root is presumably **haram-*. The *?* that precedes *m* is an allomorph of *-d*, a middle voice (MID) formative in Sidaama. There is a regular glottal metathesis rule whereby a stem-final sonorant followed by *-d* results in *?* + sonorant clusters. For instance, *fa?n-a* 'to open for oneself' comes from *fan-d-a*: *fan-* 'open', *-d* 'MID', *-a* 'infinitive'.

DOI 10.15460/auue.2025.98.1.370

2.3 Plurative

The plurative marks more than one referent. Sidaama has a wide-ranging plurative system since it has five plurative suffixes, while related HEC languages have fewer. For instance, Kambaata has two (Treis 2008: 135), while Alaaba (Schneider-Blum 2007: 67) and Qabeena (Crass 2005: 70) have three plurative formatives each. Dawit (2019: 109) claims that four "plural" markers exist in Gede'o. Appleyard (2012: 204) correctly summarizes the formation of pluratives in Cushitic languages and writes: "The formation of noun plurals is very diverse, even within groups of closely related languages […]." Although Sidaama nouns and adjectives share at least two pluratives, they also have separate plurative morphemes. Hence, the pluratives of nouns and adjectives will be discussed separately.

2.3.1 Plurative of nouns

Five pluratives of nouns are attested in the language. For ease of presentation, the pluratives of nouns (PLN) are abbreviated and numbered from one to five as: PLN1, PLN2 … PLN5.

(i) Plurative 1 of nouns

This is the most widespread plurative in Sidaama. Thus, PLN1 is formed by reduplicating the final consonant of the nominal root and by the suffixation of *-a*.

Table 11. Plurative 1 of nouns

Basic form	PLN1	Gloss
ʔakaak-o[14]	*ʔakaak-k-a*	'ancestors'
ʔawur-a	*ʔawur-r-a*	'great-grandparents'
daallas-a	*daallas-s-a*	'beds'
doog-o	*doog-g-a*	'roads'
farad-o	*farad-d-a*	'horses'
galad-o	*galad-d-a*	'monkeys'
gereb-o	*gereeb-b-a*	'sheep' (PL)
min-e	*min-n-a*	'houses'
siib-o	*siib-b-a*	'ropes'
tutum-a	*tutum-m-a*	'stumps'

14 The word *ʔakaak-o* has also the meaning 'great-grandfather'.

DOI 10.15460/auue.2025.98.1.370

Most plurative nouns in Sidaama are marked for feminine gender without consideration of the gender of the basic form and, as such, require feminine subject agreement. For instance, *min-e* 'house' is masculine while its plurative *min-n-a* 'houses' is feminine. This is seen in the gender of the modifiers that precede the nouns and the gender of the copula (see ex. 6a–6b).

(6a) *kun-i* *min-i* *hala?lado-ho*
PROX.M-NOM house-NOM wide-COP.M
'This house is wide.'

(6b) *tin-i* *min-na* *hala?l-ad-da-te*
PROX.F-NOM house-PLN1 be wide-ADJ-PLA2-COP.F
'These houses are wide.'

In (6a), the masculine noun *min-e* 'house' is preceded by the masculine proximal demonstrative *kun-i*. On the other hand, although the plurative *min-na* 'houses' is based on the masculine noun *mine* 'house', in its plurative form, it assumes a feminine gender switch, as is the case for most masculine nouns of Sidaama. That explains why the modifier that precedes *min-na* 'houses' is the feminine proximal demonstrative *tin-i*. In (6a), the basic form *mine* 'house' being a masculine subject requires the masculine copula *-ho*. On the other hand, its plurative form, i.e., *min-n-a* 'houses' in (6b), is feminine since most pluratives in Sidaama are feminine irrespective of the gender of the basic form. Hence, the copula at the end of the sentence is the feminine copula *-te*.

Although most plurative nouns in Sidaama have a feminine gender, there is at least one exception. The plurative *seenn-e* 'unmarried girls' is masculine, although its basic form, i.e., *seemm-o* 'unmarried girl/virgin', is feminine. Thus, the plurative *seenn-e* 'unmarried girls' does not obey the "feminine only" rule for pluratives. For instance, modifiers that precede *seen-e* 'unmarried girls' and the copula at the end of the clause are masculine, as shown in ex. (7).

(7) *kun-i* *seenn-i* *busule-ho*
PROX.M-NOM girls-NOM clever-COP.M
'These girls are clever.' (lit. 'this (M) girls')

In example (7), although the plurative *seenn-e* 'unmarried girls' is expected to be feminine because pluratives are mostly feminine, the

DOI 10.15460/auue.2025.98.1.370

proximal demonstrative that precedes it is *kun-i*, which is a masculine demonstrative instead of the expected feminine demonstrative *tin-i*. In addition, the copula at the end of the clause is the masculine copula *-ho* instead of the expected feminine copula *-te*.

The plurative formation in Table 11 is based on synchronic roots. However, we have also seen nouns in section 2.2.1 that have reconstructed roots. The plurative formation rule PLN 1 also applies to most of such nouns.

Table 12. Plurative 1 of diachronic roots

Diachronic root	PLN1	Gloss
**gulub-*	*gulub-b-a*	'knees'
**rumud-*	*rumud-d-a*	'roots'

Kawachi (2007: 350) posits *-adda* and *-udda* as "plural allomorphs" for the words he lists as *far-adda* 'horses', *gal-adda* 'monkeys', and *rum-udda* 'roots'. However, no such allomorphs exist in Sidaama since the first plural formation rule of nouns, i.e., PLN 1, can easily handle these words. According to my analysis they form their plurative by reduplicating the final consonant of the nominal root and by attaching *-a*. In addition, the first two segments *-ad* or *-ud* in the allomorph he proposed are the last two segments of the nominal root, as seen clearly in Table 12.

Although some of the basic forms in Table 11 are semantically plural they require a singular agreement. For instance, the basic form *farad-o* 'horses' is semantically plural but has a collective reading and hence requires a singular agreement. On the other hand, *farad-d-a* 'horses' is plurative and hence requires a plural agreement. The verbal inflections in (8b) illustrate this fact. The suffix *-tú* in (8b) marks a 3PL.PRF and a 3F.SG.PRF. However, for examples that involve pluratives, only the glossing 3PL.PRF will be provided.

(8a) *beero, farad-u sonsoom-í*
yesterday horse-M.NOM trot-3M.SG.PERF
'Yesterday, the horses (collective) trotted.' (lit. 'He trotted.')

(8b) *beero, farad-da sonsoon-tú*
yesterday horse-PL trot-3PL.PERF
'Yesterday, the horses trotted.'

DOI 10.15460/auue.2025.98.1.370

As seen in (8a), the subject is marked by *-í*, which is a marker of 3M.SG.PFV. This is because the sentence's subject is the basic form *farad-o* 'horses', which is treated as a single entity (collective) in Sidaama, although semantically, it is plural. In addition, the case suffix *-u* marks mainly the nominative of singular masculine referents.

On the other hand, in (8b), the subject is the plurative *farad-d-a* 'horses' that is treated as referring to two or more entities, and hence, this is reflected by the inflectional suffix *-tú* that marks a 3PL.PERF. In addition, there is no nominative case marking on the plurative *farad-d-a* 'horses'. This is because pluratives in Sidaama are feminine, and feminine nouns are not marked for nominative case in the language.

A set of nouns is marked chiefly by the regular singulative formative *-ʧʧo* and, in a few cases, by *-ʧo*. All of these nouns have nominal roots and the plurative of these nouns, too, is based on PLN1, as illustrated below.

Table 13. Plurative 1 of singulatives

Nominal root	Singulative	PLN1
**ʔawul-*	*ʔawul-ʧ-o* 'owl'	*ʔawul-l-a* 'owls'
**kin-*	*kin-ʧ-o* 'stone'	*kin-n-a* 'stones'
**mik'-*	*mik'-i-ʧʧ-o* 'bone'	*mik'-k'-a* 'bones'
**ʧ'ulunk'-*	*ʧ'ulunk'-i- ʧʧ-o* 'fingernail'	*ʧ'ulunk'-a* 'fingernails'
**ʧ'uuntʃ'-*	*ʧ'uuntʃ'-i- ʧʧ-o* 'big black ant'	*ʧ'uuntʃ'-a* 'big black ants'

In the examples in Table 13, PLN1 is derived by copying the final consonant of the nominal root and then by suffixing *-a*. However, in the last two nominal roots, i.e., *ʧ'ulunk'-* and *ʧ'uuntʃ'-*, reduplication of a final consonant is blocked. These two nominal roots end in consonant clusters, not single consonants. In the phonological structure of Sidaama, consonant clusters and geminates are indivisible units. Hence, a phonological rule (in this case reduplication) cannot operate on a single member of a consonant cluster or a geminate.

(ii) Plurative 2 of nouns

The formative *-aan-ʧ-o* is a productive agentive suffix that derives deverbal agentive nouns from simple verbal roots. The formative comprises the nominalizer element *-aan*, the singulative *-ʧ*, and the terminal vowel *-o*. A deverbal agentive noun usually denotes a

DOI 10.15460/auue.2025.98.1.370

habitual activity and sometimes the profession of the agent noun. The plurative of deverbal agentive nouns is formed by attaching *-o* to the nominalizer suffix *-aan* which is devoid of the singulative and the terminal vowel (see Table 14).

Table 14. Plurative 2 of deverbal agentive nouns

Simple verbal root	Singulative agentive noun	PLN2
daddal- 'trade'	*daddal-aan-ʧ-o* 'merchant'	*daddal-aan-o*[15]
dimb- 'be drunk'	*dimb-aan-ʧ-o* 'drunkard'	*dimb-aan-o*
gaʃʃ- 'rule'	*gaʃʃ-aan-ʧ-o* 'ruler'	*gaʃʃ-aan-o*
kap'p'- 'tell a lie'	*kap'p'-aan-ʧ-o* 'liar'	*kap'p'-aan-o*
kiil- 'perform sorcery'	*kiil-aan-ʧ-o* 'sorcerer'	*kiil-aan-o*
moor- 'steal'	*moor-aan-ʧ-o* 'thief'	*moor-aan-o*
nabbab- 'read'	*nabbab-aan-ʧ-o* 'reader'	*nabbab-aan-o*
ros- 'learn'	*ros-aan-ʧ-o* 'student'	*ros-aan-o*
waaʤʤ- 'fear (v.)'	*waaʤʤ-aan-ʧ-o* 'coward'	*waaʤʤ-aan-o*

Regarding the plurative of agentive nouns, Anbessa (2000: 39) incorrectly claims that "[s]ome of the agent nouns which end in /**-aančo**/ take the plural /**-aano**/" (bolding in the original source). This is because the plurative is *-o* and not *-aano*, as seen above. The formative *-aan* is a nominalizer in singulative and plurative forms.

Sidaama nouns and adjectives share many grammatical features, and sometimes the distinction between them is blurred. Thus, there is a vacillation whether some of the derived words in Table 14, such as *dimb-aan-ʧ-o* 'drunkard', *kap'p'-aan-ʧ-o* 'liar', *moor-aan-ʧ-o* 'thief', etc., can be considered as adjectives at least semantically. However, meaning alone cannot be a good criterion. In addition, although nouns and adjectives share grammatical features, they also have distinct grammatical properties. For reasons of brevity, I will not go into a detailed discussion, and I think further research may be needed to settle the issue.

Agentive nouns are formed from simple and derived verbal stems, particularly causative verbs that attach *-iis* ~ *-iʃ* ~ *-siis*. In these nouns too, the plurative formative is *-o* (Table 15). In the examples

15 The word *daddal-aan-ʧ-o* 'merchant', in addition to the regular plurative *-o*, has the plurative formative *-aasine* as in *daddal-aasine* 'merchants'.

DOI 10.15460/auue.2025.98.1.370

under Table 15 below, the basic verb stems are *ʔegen-* 'know', *ros-* 'learn', and *ʃik'-* 'approach'.

Table 15. Plurative 2 of causative agentive nouns

Derived verbal stem	Agentive noun	PLN2
ʔegen-siis- 'introduce'	*ʔegen-siis-aan-ʧ-o* 'introducer'	*egen-siis-aan-o*
ros-iis- 'teach'	*ros-iis-aan-ʧ-o* 'teacher'	*ros-iis-aan-o*
ʃik'-iʃ- 'present'	*ʃik'-iʃ-aan-ʧ-o* 'presenter'	*ʃik'-iʃ-aan-o*[16]

(iii) Plurative 3 of nouns

The third plurative in Sidaama is PLN3, formed by suffixing *-uba* to the final consonant of the nominal root. Although the list contains disparate nouns, some plural nouns formed via PLN3 are basic forms related to kinship terms or paired body parts. In Sidaama /b/ becomes [*w*] intervocalically. Hence in all the pluratives listed below, the phonetic form of the plurative will become [*-uwa*] as in: [*ʔadd-uwa*] 'heifers', [*ball-uwa*] 'in-laws', [*lekk-uwa*] 'legs', etc.

Table 16. Plurative 3 of nouns

Basic form	PLN3	
ʔadd-e	*ʔadd-uba*	'heifers'
ʔam-a	*ʔam-uba*	'mothers'
ʔann-a	*ʔann-uba*	'fathers'
ʔill-e	*ʔill-uba*	'eyes'
ʔamat't'-o	*ʔamat't'-uba*	'spears'
ball-o	*ball-uba*	'in-laws'
bass-a	*bass-uba*	'scars'
lekk-a	*lekk-uba*	'legs'
matʃ'ʧ'-a	*matʃ'ʧ'-uba*	'ears'
su'm-a	*su'm-uba*	'names'
rod-oo	*rod-uuba*	'brothers/sisters'
ʔar-oo	*ʔar-uuba*	'husbands'

16 It appears in Shimelis (2007: 117) as part of the phrase *duduwu shiqishaancho*, 'news presenter'.

DOI 10.15460/auue.2025.98.1.370

It seems that basic forms with long terminal vowels will also have long vowels in their pluratives, as the last two examples in Table 16, i.e. *rod-uuba* 'brothers/sisters' and *ʔar-uuba* 'husbands' demonstrate.

(iv) Plurative 4 of nouns

The formative of the fourth plurative, i.e., PLN4, is formed by suffixing *-ubba* to the final consonant of the nominal root. As can be seen in Table 17, most of the nouns that are pluralized via PLN4 are basic forms related to paired body parts. Note that three of the paired body parts, i.e., *ʔill-e* 'eye', *lekk-a* 'leg', and *matʃ'tʃ'-a* 'ear' can also be pluralized via the formative *-uba*, as seen in Table 17. This indicates that some paired body parts have alternative pluratives.

Table 17. Plurative 4 of nouns

Basic form	PLN4	
ʔang-a	*ʔang-ubba*	'hands'
ʔill-e	*ʔill-ubba*	'eyes'
buud-a	*buud-ubba*	'horns'
gordo	*gord-ubba*	'heavens'
lekk-a	*lekk-ubba*	'legs'
matʃ'tʃ'-a	*matʃ'tʃ'-ubba*	'ears'
sik'k'-o	*sik'k'-ubba*	'sticks'
ʃomb-o	*ʃomb-ubba*	'lungs'

The formative of PLN4, i.e., *-ubba*, is similar to *-uba*, a plurative of PLN3. In PLN3, the segment *b* is single; in PLN4, it is *bb*, i.e., a geminate. Because of this reason, Hudson (1976: 252) writes, "[…] this suggests a reconstructed plural suffix uba which ordinarily gave uwa except where stems with b encouraged retention of the b of the suffix, which was then geminated" (underlining in the original). However, Hudson's claim has two drawbacks. First, many nouns do not have an internal *b* but take the plurative *-ubba* as most of the examples in Table 17 illustrate. In addition, at least two nouns in Table 16 contain a stem-internal *b* but their plurative formative is *-uba* and not the expected *-ubba*. The two nouns are *ball-o*, 'in-law', and *bass-a,* 'scar', whose pluratives are *ball-uba*, 'in-laws' and *bass-uba*, 'scars', respectively.

DOI 10.15460/auue.2025.98.1.370

(v) Plurative 5 of nouns

Plurative 5 (PLN5) is formed by suffixing *-oota* to the basic form which is devoid of a terminal vowel. These nouns have all the number categories of Sidaama: basic form, singulative, and plurative (Table 18).

Table 18. Plurative 5 of nouns

Basic form	Singulative	PLN5
harr-e 'donkeys'	*harr-i-ʧʧ-o* 'donkey'	*harr-oota* 'donkeys'
mann-a 'men/ people'	*man-ʧ-o* 'person'	*mann-oota* 'people/men'
woʃʃ-a 'dogs'	*woʃ-i-ʧʧ-o* 'dog'	*woʃʃ-oota* 'dogs'

As we shall see later on, the above three basic forms belong to the category of inherently plural nouns which refer to a collective or group. Although these nouns are semantically plural, they are treated in Sidaama as a single entity because they refer to a collective or group. Two grammatical properties of such nouns that refer to a collective are: they require a singular morphological agreement and a masculine predicative copula *-ho*. On the other hand, nouns such as *mann-oota* 'people/men' and *woʃʃ-oota* 'dogs' (which are based on PLN5) are ordinary plurals because they refer to two or more referents. The evidence again comes from verbal inflections, as shown in (9a–9b).

(9a) *man-n-u* *sirb-í*
man-PL-NOM sing-3M.SG.PERF
'The people/men sang.'

(9b) *mann-oot-u* *sirb-i-tú*
man-PL-NOM sing-3PL.PERF
'The people/men sang.'

In (9a), the basic form *man-n-a* 'people/men', is considered a single entity and hence is marked by *-í*, a singular inflection that marks a 3M.SG.PERF. On the other hand, in (9b), the plurative *mann-oota* 'people/men' is based on PLN5, which is an ordinary plural, and that is the reason why it is marked by a plural inflection, i.e., *-tú* which marks a 3PL.PERF.

2.3.2 Pluratives of adjectives

Adjectives can be pluralized like nouns, and the number of adjective pluratives is similar to that of nouns. Nevertheless, some pluratives

DOI 10.15460/auue.2025.98.1.370

are shared both by nouns and adjectives. For instance, PLN5 of nouns based on the formative *-oota* is identical to the adjectival plurative PLA5. In addition, PLN1 of nouns is identical to PLA1 of adjectives. Both are formed similarly: by reduplicating the final consonant of a nominal root or an adjectivizer stem and then by suffixing *-a*.

(i) Plurative 1 of adjectives

The suffix *-aam-o* is a productive morpheme that derives denominal adjectives from nouns[17] and has the meaning of 'full of N' or 'characterized by N' (where N = noun). The formative *-aam-o* is suffixed to a nominal root, i.e., a form devoid of the terminal vowel. Denominal adjectives in Sidaama distinguish masculine and feminine gender with the suffixation of *-o* and *-e*, respectively, as in: *matʃtʃaraar-aam-o* 'mad/crazy (M)' and *matʃtʃaraar-aam-e* 'mad/crazy (F)'.

The first plurative of adjectives, i.e., PLA1, is formed by reduplicating *m*, i.e., the final consonant of the adjective stem *-aam*, and then by suffixing *-a*. Consider the examples in Table 19 where the adjectives are in masculine gender.

Table 19. Plurative 1 of adjectives

Noun	Adjective	PLA 1
ʔagood-a 'shoulder'	*ʔagood-aam-o* 'muscular'	*ʔagood-aam-m-a*
ʔamal-e 'temper'	*ʔamal-aam-o* 'ill-tempered'	*ʔamal-aam-m-a*
ʔatoot-e 'riches'	*ʔatoot-aam-o* 'full of riches'	*ʔatoot-aam-m-a*
ʔawuub-e 'white hair'	*ʔawuub- aam-o* 'full of white hair'	*ʔawuub- aam-m-a*
badd-e 'bald spot'	*badd-aam-o* 'bald'	*badd-aam-m-a*
biʤaaʤ-o 'scabies'	*biʤaaʤ-aam-o* 'scabby'	*biʤaaʤ-aam-m-a*
buluul-o 'ashes'	*buluul-aam-o* 'covered with ash'	*buluul-aam-m-a*
buud-a 'horn'	*buud-aam-o* 'with horns'	*buud-aam-m-m-a*
matʃtʃaraar-o 'madness'	*matʃtʃaraar-aam-o* 'mad/crazy'	*matʃtʃaraar-aam-m-a*

17 I came across one example whereby an adjective is derived from a verb. Thus the verb *ʔafiɗ-* 'get' yields the adjective *ʔafiʔr-aam-o* 'wealthy' and the plural *ʔafiʔr-aam-m-a*.

DOI 10.15460/auue.2025.98.1.370

(ii) Plurative 2 of adjectives

Many adjectives are pluralized with the suffix *-ulle*. There are seven semantic classes of adjectival concepts that were proposed by Dixon (1982: 16). Amongst these, six of the semantic classes contain 12 adjectives in Sidaama that are pluralized based on *-uulle* i.e., PLA2, which is attached to the adjectival root i.e. a form that is devoid of the terminal vowel. The only semantic field that was not involved is SPEED since there are no adjectives in Sidaama for its expression. A similar phenomenon is observed in Alaaba (Schneider-Blum 2007: 118) and Kambaata (Treis 2008: 254). Table 20 provides examples for PLA2 based on various semantic fields.

Table 20. Plural 2 of adjectives

Semantic field	Adjective	PLA 2
DIMENSION	*duʔm-a* 'fat'	*duʔm-uulle*
	seed-a 'tall'	*seed-uulle*
PHYSICAL PROPERTY	*faayy-a* 'beautiful'	*faayy-uulle*
	t'um-a[18] 'good'	*t'um-uulle*
COLOR	*biiʃʃ-a* 'red'[19]	*biiʃʃ-uulle*
	waaʤʤ-o 'white'	*waaʤʤ-uulle*
	haanʤ-a 'green'	*haanʤ-uulle*
HUMAN PROPENSITY	*goww-a* 'foolish'	*goww-uulle*
AGE	*ʔaʤ-a* 'young'	*ʔaʤ-uulle*
	maaʔn-e 'younger in age'	*maaʔn-uulle*
VALUE	*buʃ-a* 'bad'	*buʃ-uulle*
	dantʃ-a 'good'	*dantʃ-uulle*

(iii) Plurative 3 of adjectives

Few adjectives are formed by suffixing *-ad-o* to pre-categorial roots, whereby *-ad* is the adjectivizer formative, while *-o* is a terminal vowel. The plurative of this group of adjectives is formed by reduplicating the final consonant of the adjectivizer formative, i.e., *d*, and then by suffixing *-a*, as shown in Table 21 below.

18 Also has the meaning of beautiful.

19 A color for human beings with light-colored hue.

DOI 10.15460/auue.2025.98.1.370

Table 21. Plurative 3 of adjectives

Pre-categorial root	Adjective	PLA 3
**ʔairr-* 'become honored'[20]	*ʔairr-ad-o* 'honored'	*ʔairr-ad-d-a*
**ʔiibb-* 'become hot'	*ʔiibb-ad-o* 'hot'	*ʔiibb-ad-d-a*
**kaaʗʗ-* 'become thin'	*kaaʗʗ-ad-o* 'strong'	*kaaʗʗ-ad-d-a*
**ʃaal-* 'be slim'	*ʃaal-ad-o* 'thin'	*ʃaal-ad-d-a*
**ʃak'k'-* 'be soft'	*ʃak'k'-ad-o* 'soft'	*ʃak'k'-ad-d-a*

(iv) Plurative 4 of adjectives

These pluratives are formed by reduplicating the final consonant of a pre-categorial root and then suffixing the plurative *-aadda*, as shown in Table 22 below.

Table 22. Plurative 4 of adjectives

Pre-categorial root	Adjective	PLA 4
**busul-*	*busul-e* 'clever'	*busul-l-aadda*
**duum-*	*duum-o*[21] 'red'	*duum-m-aadda*
**ʃiim-*	*ʃiim-a* 'small'	*ʃiim-m-aadda*
**bair-*	*bair-a* 'older sibiling'	*bair-r-aadda*
**ʗaaw-*	*ʗaaw-aʃ-ʃ-o* 'thin'	*ʗaaw-aadda*

As can be seen in Table 22, in the last pre-categorial i.e., *ʗaaw-*, its last consonant i.e. *w* does not undergo reduplication.

(v) Plurative 5 of adjectives

These pluratives are formed by suffixing the formative *-oota* to pre-categorial roots. Thus, PLA5 is identical to PLN5 of nouns that contain the same suffix.

Table 23. Plurative 5 of adjectives

Pre-categorial root	Adjective	PLA 5
**babb-* 'stammer'	*babb-i-ʧʧ-a* 'stammerer'	*babb-oota*
**ball-* 'be blind'	*ball-i-ʧʧ-a* 'blind'	*ball-oota*
**dank'-* 'be deaf'	*dank'-i-ʧʧ-a* 'deaf'	*dank'-oota*
**murk'-* 'be cut off'	*murk'-i-ʧʧ-a* 'dismembered'	*murk'-oota*

20 The pre-categorial root **ʔairr-* has also the meaning of 'become heavy' and 'become pregnant'.

21 This adjective exhibits a masculine feminine distinction: *duum-o* 'red (M)' and *duum-e* 'red (F)'.

DOI 10.15460/auue.2025.98.1.370

Pre-categorial root	Adjective	PLA 5
**doofaar-* ‘be untidy’	*doofaar-a* ‘untidy’	*doofaar-oota*
**goww-* ‘be foolish’	*goww-a* ‘foolish’	*goww-oota*

In examples in Table (23), the last four adjectives distinguish gender, whereby *-ʧʧ-a* marks masculine, and *-tt-e* marks feminine as in *ball-i-ʧʧ-a* ‘blind (m)’ and *ball-i-tt-e* ‘blind (f).’ Since *-ʧʧ* also marks a singulative, hence *-ʧʧa* can be considered as a portmanteau morpheme that marks two functions: gender and number.

(vi) Plurative 6 of adjectives

Several adjectives are derived from pre-categorial roots with the suffixation of two formatives that distinguish gender: *-all-eessa* (M) and *-all-eette* (F) as in *ʧ’ee?m-all-eessa* ‘lazy (M)’ and *ʧ’ee?m-all-eette* ‘lazy (F)’. The function of the element *-all* in the adjectivizing formative of Sidaama is unclear. The same holds true for related HEC languages Kambaata and Qabeena. Regarding Kambaata Treis (2008: 283) writes, “[t]he morpheme *-all* might be a former derivational morpheme having served to derive adjectives from verbs. In modern Kambaata, it is as unproductive as in Qabeena […]”. According to Crass (2005: 73) in Qabeena, the function of the suffix *-all* cannot be determined.[22]

The elements *-eessa* and *-eette* in the adjectivizing formative are portmanteau morphemes that mark gender on one hand and number (singulative) on the other. The plurative of these adjectives is based on *-eeyye*, a marker of PLA6. For sake of brevity, the adjectives will be based only on the masculine form: *-all-eessa* (Table 24).

Table 24. Plurative 6 of adjectives

Pre-categorial root	Adjective	PLA 6
**boon-* ‘boast’	*boon-all-eessa* ‘boastful’	*boon-all-eeyye*
**ʧ’ee?m-* ‘be lazy’	*ʧ’ee?m-all-eessa* ‘lazy’	*ʧ’ee?m-all-eeyye*
**fayy-* ‘be healthy’	*fayy-aall-eessa* ‘healthy’	*fayy-aall-eeyye*
**waʤʤ-* ‘be afraid’	*waʤʤ-all-eessa* ‘coward’	*waʤʤ-all-eeyye*
**?af-* ‘know’	*?af-all-eessa* ‘wise man’	*?af-all-eeyye*

22 German-English translation of the relevant sentence was courtesy of Ronny Meyer to whom I am indebted.

DOI 10.15460/auue.2025.98.1.370

There are also a few other adjectives whose plurative is based on PLA 6, i.e., *-eeyye*, which can be added as a subset of the above adjectives. They are listed in Table 25 below.

Table 25. Additional adjectives based on plurative 6

Root	Adjective	PLA 6
**ʔog-* 'be wise'	*ʔog-eessa* 'bonesetter, professional'	*ʔog-eeyye*
**mah-* 'be last'	*mah-eessa* 'last in a group'	*mah-eeyye*
**soor-* 'lead'	*soor-eessa* 'first, preceding'	*soor-eeyye*
**ʧ'im-* 'arbitrate'	*ʧ'im-eessa* 'elder of a clan, arbiter'	*ʧ'im-eeyye*

The adjectives in Table 25, too, make gender distinctions through the same formatives listed in Table 24: *-eessa* (M) and *-eette* (F). Thus, *ʔog-eessa* is 'bonesetter, professional' while according to Gasparini (1983: 249), *ʔog-eette* is 'midwife'. Among this group, only the last word, *ʧ'im-eessa*, does not have a feminine counterpart, probably because of cultural norms since only males are nominated to be arbiters or elders of their clan.

Two differences exist between the adjectives listed in Table 24 and those in Table 25. Adjectives in Table 25 lack the obscure linguistic element *-all* which is found in adjectives in Table 24. In addition, adjectives in Table 24 are derived from pre-categorial roots that allow the derivation of different word classes, while the roots in Table 25 allow adjectival derivation only. For instance, from the pre-categorial root *ʧ'ee?m-* 'be lazy' in Table 24, in addition to the adjective *ʧ'ee?m-all-eessa* 'lazy', it is also possible to derive a noun *ʧ'ee?m-ile* 'laziness' (via the nominalizer *-le*) and a verb *ʧ'ee?m-í* 'he became lazy' (by suffixing *-í*, a 3M.SG.PERF).

3 Functions of number marking morphemes in Sidaama

In the second section of this paper, the number-marking morphemes of Sidaama were dealt with and this included a detailed discussion of singulatives and pluratives of nouns and adjectives. In this chapter, the functions of singulatives and pluratives will be discussed.

3.1 Functions of singulative

In Sidaama, the singulative marks a single referent. Nonetheless, the singulative formative is not found on all nouns that denote a single

DOI 10.15460/auue.2025.98.1.370

referent. For instance, basic forms such as *min-e* 'house', *doog-o* 'road', *sik'k'-o* 'stick', etc. denote a single referent but do not carry a singulative formative. According to Treis (2008: 137) "[…] there is a class of nouns, predominantly referring to humans and higher animals, which are already inherently singular.".

(i) Inherently singular nouns

In Sidaama, inherently singular nouns are devoid of a singulative marker, and hence, their basic form is employed to denote a single referent. Table 26 presents a list of inherently singular nouns of Sidaama, similar to that of Kambaata (Treis 2008: 137; 2014: 5).

Table 26. List of inherently singular nouns of Sidaama

ʔam-a 'mother'	*ʔadd-e* 'heifer'
ʔann-a 'father'	*boot-o* 'young bull'
ball-o 'in-law'	*gaang-o* 'mule'
beett-o[23] 'boy/girl'	*gaŋŋ-a* 'mare'
ʤaal-a 'friend'	*hand-o* 'ox'
k'aak'k'-o 'infant'	*saa* 'cow'
rod-oo 'sibling'	*siito* 'calf'
seemm-o 'virgin girl'	

With a number that exceeds one, the inherently singular noun itself or its plurative equivalent is used. In both cases, the numeral precedes the plurative or the inherently singular noun, as shown in ex. (10a) and (10b).

(10a) *kaayyamo* *sas-e* *hand-a* *hir-í*
Kaayyamo three-ACC ox-PLN1 sell-3M.SG.PERF
'Kaayyamo sold three oxen.'

(10b) *kaayyamo* *sas-e* *hand-o* *hir-í*
Kaayyamo three-ACC ox-ACC sell-3M.SG.PERF
'Kaayyamo sold three oxen.' (lit. 'three ox')

As can be seen in in the first sentence (10a), the plurative *hand-a* 'oxen' is used, while in the second sentence (10b), the basic form of the inherently singular noun, i.e., *hand-o* 'ox (ACC)', has replaced it.

23 Sidaama does not allow the suffixation of a singular to inherently singular nouns. The only exception appears to be *beetto* 'boy/girl'. With the suffixation of the singulative *ʧʧ-o* the resultant noun is *beett-i-ʧʧ-o* 'a small girl'.

DOI 10.15460/auue.2025.98.1.370

In both sentences, the numeral *sas-e* 'three' precedes the plurative or the inherently singular noun. This illustrates that number agreement with a head noun is not obligatory in Sidaama. Nevertheless, a numeral with a plurative is the preferred option.

Inherently singular nouns are distinct from singulative nouns in that an overt singulative marker does not mark them. Nonetheless, an inherently singular noun denotes a single referent like a singulative; hence, both behave similarly grammatically, as the example below illustrates.

(11) *hand-u* *woš-i-ʧʧ-o* *ʃorr-í*
ox-NOM.M dog-EE-SG-ACC chase-3M.SG.PERF
'The ox chased the dog.'

(ii) Singulative as a marker of diminutive

Singulative is hardly marked on adjectives save for very few ones, such as *but'-i-ʧʧ-o* 'poor' or *ʤaaw-aʃ-ʃ-o* 'thin'. However, there are three adjectives in Sidaama that suffix a special singulative to express diminutive. The singulative allomorph that is involved in such cases is not the regular singulative *-ʧʧ-o* but rather *-ʧ'ʧ'-o* whereby the ʧ element is glottalized. The diminutive is formed by suffixing *-ʧ'ʧ'-o* to the adjectival stem devoid of the terminal vowel. An epenthetic *-i* is inserted to break impermissible consonant clusters in all the three diminutives below. In such cases, the resultant adjective denotes feminine referents only.

Table 27. Singulative as a marker of diminutive

Adjective	Adjectives with the singulative *-ʧ'ʧ'-o*
ʃiim-a 'small'	*ʃiim-i-ʧ'ʧ'-o* 'a very small cute girl'
dantʃ-a 'good, beautiful'	*dantʃ-i-ʧ'ʧ'-o* 'a small good-looking girl'
faayy-a 'pretty'	*faayy-i-ʧ'ʧ'-o* 'a small pretty girl'

(iii) Singulative as a marker of a small amount

Sidaama contains several mass nouns such as *mat'in-e* 'salt', *bull-ee* 'flour', *waa* 'water', *buuro* 'butter', *farʃ-o* 'local beer', *maal-a* 'meat', etc. These mass nouns being non-count nouns necessitate the use of traditional measure nouns such as *safe* 'a grain measure', *hoowo* 'palmful' or modern measure nouns (which are exclusively loanwords) such as *kilo* 'kilogram' or *liitiro* 'liter'. Mass nouns do not take pluratives. However, three mass nouns can suffix a singulative to express a notion of 'piece of MN' or 'a small amount of MN', whereby

DOI 10.15460/auue.2025.98.1.370

MN stands for a mass noun as illustrated below. The singulative is attached to the nominal stem i.e., a form devoid of a terminal vowel.

Table 28. Singulative as a marker of small amount

Mass noun	Singulative of mass noun
bull-ee 'flour'	*bull-ii-ʧʧ-o* 'a small amount of flour'
buur-o 'butter'	*buur-ʧ-o* 'a small amount of butter'
maal-a 'meat'	*maal-ʧ-o* 'a piece of meat'

Few singulatives and inherently singular nouns exhibit a 'gender switch' whereby feminine nouns in the singular or singulative can trigger masculine agreement in the plural as was claimed by Kramer & Anbessa (2020: 1).

(12a) *tin-i* *seemm-o*
PROX.F-NOM unmarried girl
'this unmarried girl'

(12b) *kun-i* *seenn-i*
PROX.M-NOM unmarried girlS-NOM
'these unmarried girls' (lit. 'this' (M)'

(13a) *tin-i* *woʃ-i-ʧʧo*
PROX.F-NOM dog-EE-SG
'this female dog (bitch)'

(13b) *kun-i* *woʃʃ-oot-i*[24]
PROX.M-NOM bitches-PLN4-NOM
'these bitches/dogs' (lit. 'this (M)'

In (12a) in the phrase *tin-i seemmo* 'this unmarried girl', the proximal demonstrative *tin-i* 'this' is marked for feminine because the head noun *seemmo* 'unmarried girl' is feminine in gender. However, in (12b), the gender marking of the proximal demonstrative is switched from feminine to masculine: from *tin-i* 'this (F)' to *kun-i* 'this (M)' despite *seenn-e* 'unmarried girls' which is the plurative equivalent of *seemm-o* 'unmarried girl' must be treated as feminine in gender. Note that in Sidaama, most pluratives are treated as feminine. In relation

24 In this phrase, the head noun can also be *woʃʃ-i* 'dogs' (NOM), an inherently plural noun.

DOI 10.15460/auue.2025.98.1.370

to this, Kramer & Anbessa (2020: 7) write, "[i]n the plural, almost all Sidaama nouns trigger feminine agreement regardless of what gender they have in the singular".

3.2 Functions of plurative

Sidaama has a list of inherently plural nouns that designate a collective or a group.[25] In such nouns, the basic form marks an inherently plural noun, similar to an inherently singular noun.

In inherently plural nouns, a singulative form is necessary to denote a singular referent. Again, the list of inherently plural nouns of Sidaama is similar to that of Kambaata (Treis 2008: 138; 2014: 5).

Table 29. List of inherently plural nouns of Sidaama

fiit't'-a 'relatives'	*farad-o* 'horses'
geerr-a 'elders'	*galad-o* 'monkeys'
mann-a 'men, people'	*gereeb-o* 'sheep' (PL)
meent-o 'women'	*got-o* 'hyenas'
seenn-e 'unmarried girls'	*hamas-o* 'snakes'
wedell-a 'youngsters'	*me?-e* 'goats'
Sidaam-u 'Sidaama people'	*woʃ-ʃ-a* 'dogs'
wosin-a 'guests'	

Most inherently plural nouns listed above in Table 29 have a corresponding singulative. For instance, *geer-ʧo* 'elder', *faraʃ-ʃo* 'horse', and *me?-i-ʧʧo* 'goat' are singulatives of *gerr-a* 'elders', *farad-da* 'horses', and *me?-?a* 'goats', respectively. The only exception without a corresponding singulative appears to be *seen-e* 'unmarried girls'. Even in this case, *seenne* has a basic form *seem-o* 'unmarried girl', which is inherently singular and hence refers to a single referent. The inherently plural nouns *mann-a* 'men, people' and *meent-o* 'women' share the singulative *manʧ-o* 'man/woman.'

Inherently plural nouns refer to a collective of entities as a single group and are marked morphologically for singular. This includes modifiers that precede them that are marked for masculine.

(14) *mee?-u* *t'ooq-í*
goats-NOM.M flee-3M.SG.PERF
'The goats fled.' (lit. 'He fled.')

25 According to Treis (2008: 138), the same holds in Kambaata.

DOI 10.15460/auue.2025.98.1.370

(15) *kun-i* *galad-i* *ʔooso*
PROX.M-NOM.M monkeys-NOM children
ʃorr-í
chase-3M.SG.PERF
'These monkeys chased the children.' (lit. 'this monkeys')

In ex. (14), although *meʔ-e* 'goats' designates a group of at least two goats and even more, this noun and others in the list in Table 29 are treated as a collective in Sidaama. Therefore, this explains why the verb is in singular *t'ooq-í* 'he fled' instead of the expected plural *t'ooq-qú* 'they fled'. Consequently, the inflectional features on the verb are the same for inherently plural nouns and singulatives because both are treated as a single entity in Sidaama. Likewise, the masculine proximal *kun-i* 'this' in (15) that precedes the head noun *galad-i* 'monkeys' is in singular instead of *kur-i* 'these' the expected plural proximal.

Consider (16) whereby the inherently plural *galad-o* 'monkeys' of (15) is replaced by the singulative *galaʃ-ʃ-o* 'monkey'.

(16) *kun-i* *galaʃ-ʃ-i* *ʔooso*
PROX.M-NOM.M monkey-SG-NOM children
ʃorr-í
chase-3M.SG.PERF
'This monkey, chased the children.'

Although inherently plural nouns are viewed as designating a collective entity, some have plurative forms. For instance, the inherently plural *galad-o* 'monkeys' has a plurative form *galad-d-a* 'monkeys' that refers to two or more referents. Accordingly, the verbal inflection will also reflect this fact, as shown in ex. 17).

(17) *beero* *galad-da* *ʔooso* *ʃorr-i-tú*
yesterday monkey-PL children chase-EE-3PL.PERF
'Yesterday, the monkeys chased the children.'

Two inherently plural nouns are exceptional because they do not have a corresponding singulative derived from them, unlike the list in Table 29. These nouns are *ʔoos-o* 'children' and *lal-o* 'cattle'. This also holds true for Kambaata, as Treis (2008: 138) mentions. For a single referent, inherently singular nouns are used: in the case of

DOI 10.15460/auue.2025.98.1.370

ʔoos-o ‘children’ *beett-o* ‘boy, son/girl, daughter’ and for *lalo* ‘cattle’, either *hand-o* ‘ox’ or *saa* ‘cow’ will be used.

3.3 Transnumeral nouns

In addition to inherently singular and plural nouns, there are transnumeral nouns in Sidaama. These nouns can refer to a single entity or plural entities in their basic form. In Sidaama, paired body parts are transnumeral. The list is presented in Table 30.

Table 30. List of inherently transnumeral nouns

ʔagan-a ‘moon’	*dikk-o* ‘market(s)’
ʔang-a ‘hand(s)’	*hak’k’-a* ‘tree(s)’
ʔill-e ‘eye(s)’	*lekk-a* ‘leg(s)’
ʔunun-a ‘breast(s)’	*matʃ’tʃ’-a* ‘ear(s)’
dar-o ‘leaf/leaves’	

Many transnumerals have plurative forms. For instance, *ʔang-a* ‘hand’ has the plurative *ʔang-ubba* ‘hands’. Although that is the case, the basic form *ʔang-a* ‘hand’ is used in daily speech as shown under (18).

(18) *sood-o* *ʔanga-si* *haiʃʃ-i-ɗ-í* [*haiʃʃ-i-ʔr-í*]
morning hand-3M.SG.GEN wash-EE-MID-3M.SG.PERF
‘In the morning he washed his hands.’

Although *ʔang-a* ‘hand’ in (18) is a basic form and is assumed to refer to one hand, it is interpreted as ‘hands’ based on pragmatics. People typically wash both hands and not one. This implies that although a noun can refer to plural referents, a plurative form is not obligatory. This is particularly true for transnumerals since they can refer to one or more than one referent. In the case of *ʔanga* ‘hand(s)’, it has a plurative *ʔang-ubba* ‘hands’. However, replacing *ʔanga* in sentence (18) with *ʔang-ubba* will render it ungrammatical, as shown below in (19).

(19) **sood-o* *ʔang-ubba-si* *haiʃʃ-i-ɗ-í* [*haiʃʃ-i-ʔr-í*]
morning hand-PLN4-3M.SG.GEN wash-EE-MID-3M.SG.PERF
‘In the morning he washed his hands.’

4 Conclusion

This paper examined the number markings on Sidaama nouns and adjectives in detail. There are three formal categories of number in

DOI 10.15460/auue.2025.98.1.370

Sidaama: basic form, singulative, and plurative. The basic form is a noun unmarked for number; a singulative denotes a single referent, and a plurative marks more than one referent.

The singulative of nouns is marked mainly by *-ʧ-o* ~ *-ʧʧ-o*. There are also other formatives such as *-ʃ-o*, *-ʧ-o*, and *-k-o* that function as markers of the singulative synchronically. Such formatives arose from the merging of the stem-final consonant of the basic stem with the initial *-ʧ* of the singulative. Singulative is hardly marked on adjectives save for very few ones such as *ʤaawaʃ-ʃ-o* 'thin.' Five types of nominal pluratives were identified. These are: *-Ca*, *-o*, *-uba*, *-ubba*, and *-oota*. The first one, i.e., *-Ca*, is the most frequent, and it involves copying of a stem-final consonant and suffixation of *-a*. The six adjectival pluratives are: *-ma*, *-uulle*, *-da*, *-Caadda*, *-oota* and *-eeyye*. As can be seen from the plurative list, nouns and adjectives share two of the plural formatives: *-Ca* and *-oota*. During the discussion on number marking, it has become clear that the word class of most basic stems is indeterminate. This is because the same stem allows the derivation of different word classes; hence, it is difficult to determine the basic word class of the stem itself.

Sidaama contains inherently singular nouns that comprise the domain of humans and higher animals. Their basic form marks a single referent. The language also contains inherently plural nouns that refer to a collective of entities or a group. In such nouns, a singulative form is obligatory to denote a singular referent. Most of the inherently plural nouns have a corresponding singulative. Few nouns are transnumeral, and in their basic form, they can denote either a single referent or a plural referent.

Some elements necessitate future research because they cannot be explained at the present stage. For instance, there are many stems in Sidaama whose basic word class cannot be determined. Moreover, there is a question of whether *-aan-ʧ-o* can also derive agentive adjectives and not only agentive nouns. Hence, there is a plan to incorporate more written corpus and spoken data. I hope this may enable me to answer the questions raised here and to explore further number marking in Sidaama.

DOI 10.15460/auue.2025.98.1.370

List of Abbreviations

3 third person, ACC accusative, ADJ adjectivizer, COP copula, COP.F feminine copula, COP.M masculine copula, EE epenthetic element, F feminine, GEN genitive, IMPERF imperfect, INF infinitive, M masculine, MID middle voice, NEG negative, NOM nominative, PERF perfect, PL Plural; Plurative, PLA plurative of adjective, PLN plurative of noun, PROX proximal, SG singular, SGV singulative.

References

Anbessa Teferra. 2012. *A grammar of Sidaama: Phonology, morphology, and syntax.* Saarbrücken: LAP LAMBERT Academic Publishing.

Anbessa Teferra. 2014. *Sidaama (Sidaamu Afoo).* (Languages of the World/ Materials 501). München: Lincom GmbH.

Appleyard, David. 2012. Cushitic. In Lutz Edzard (ed.), *Semitic and Afroasiatic: Challenges and opportunities.* Wiesbaden: Harrassowitz. 199–292.

Azeb Amha. 2017. The Omotic language family. In Alexandra Y. Aikhenvald & Robert M.W. Dixon (eds.), *The Cambridge handbook of linguistic typology*. Cambridge: Cambridge University Press. 815–853. https://doi.org/10.1017/9781316135716.026

Central Statistical Authority. 2010. *The 2007 Population and Housing Census of Ethiopia.* Addis Ababa: Statistical Report for Southern Nations, Nationalities and Peoples' Region.

Crass, Joachim. 2005. *Das K'abeena. Deskriptive Grammatik einer hochland-ost-kuschitischen Sprache.* Köln: Rüdiger Köppe Verlag.

Dixon, R.M.W. 1982. *Where have all the adjectives gone? And other essays in semantics and syntax.* Berlin: Mouton.

Dawit Tilahun Jembere. 2019. *A grammar of Gede'o: A Highland East Cushitic language.* Addis Ababa: Addis Ababa University Ph.D dissertation.

Gasparini, Armido. 1983. *Sidamo-English dictionary.* Bologna: Editrice Missionara Italiana.

Hudson, Grover 1976. Highland East Cushitic. In Lionel M. Bender, (ed.), *The Non-Semitic languages of Ethiopia.* East Lansing: African Studies Center, Michigan State University. 232–277.

Hudson, Grover. 1989. *Highland East Cushitic dictionary*. Hamburg: Buske.

Kawachi, Kazuhiro. 2007. *A grammar of Sidaama (Sidamo), a Cushitic language of Ethiopia*. Buffalo: University of Buffalo Doctoral dissertation.

DOI 10.15460/auue.2025.98.1.370

Kramer, Ruth & Anbessa Teferra. 2020. Gender switch in Sidaama. *Brill's Journal of Afroasiatic Languages and Linguistics* 12(2). 286–327. https://doi.org/10.1163/18776930-01102006

Moreno, Martino Mario. 1940. *Manuale di Sidamo*. Milano: Mondadori.

Schneider-Blum, Gertrud 2007. *A grammar of Alaaba. A Higland East Cushitic language of Ethiopia.* (Cushitic Language Studies 25). Köln: Rüdiger Köppe Verlag.

Shimelis Gizaw (ed.). 2007. *Sidaama-Amaarunna-Inglizete qalla borraasincho (Sidaama-Amharic-English dictionary)*. Addis Ababa: Master Printing Press.

Treis, Yvonne. 2008. *A grammar of Kambaata (Part 1): Phonology, nominal-morphology, and non-verbal predication*. (Cushitic Language Studies 26). Köln: Rüdiger Köppe Verlag.

Treis, Yvonne 2014. Number in Kambaata. In Anne Storch & Gerrit J. Dimmendaal (eds.), *Number: Constructions and semantics. Case studies from Africa, India, Amazonia and Oceania.* (Studies in Language Companion Series 151). Amsterdam & Philadelphia: John Benjamins. 111–134. https://doi.org/10.1075/slcs.151.05tre

Yri, Kjell Magne. 2011. The singulative in Sidaaamu Afó. *Folia Orientalia* 47(2). 117–136.

Yri, Kjell Magne. 2012. Sidaama. In Lutz Edzard (ed.), *Semitic and Afroasiatic: Challenges and opportunities.* Wiesbaden: Harrassowitz. 262–277.

DOI 10.15460/auue.2025.98.1.370

Cushitic loans in South Nilotic revisited: A deconstruction of Proto Baz

Maarten Mous[a] *& Christian Rapold*[b]

Leiden University[ab]
m.mous@hum.leidenuniv.nl
hobasun37@gmail.com

Abstract
This article explores lexical contact between Cushitic and Nilotic taking the proposed Cushitic language Proto Baz as point of departure. Proto Baz is a putative East Cushitic language proposed by Heine, Rottland & Vossen (1979), on the basis of words borrowed into South Nilotic. Some of the loans were already pointed out by Ehret (1970, 1971, 1974). A critical assessment of the proposed loans shows that there is more diversity in their origin than formerly thought, and that it is in fact not necessary to assume the existence of a separate language (Proto Baz) to account for the loans. Instead, the loans came from various reconstructed stages of existing Cushitic languages. We discuss different contact situations between groups of East Cushitic and South Nilotic languages, as well as the recurring challenges in historical interpretation of the available evidence. The Appendix provides the reader with an overview of the discussed lexical items from earlier sources and new proposed analyses.

Keywords: Proto Baz, Cushitic, Proto South Nilotic, loan words, language contact, Kenya, Ethiopia

1 Introduction[1]

This article is a re-assessment of the proposals by Heine, Rottland & Vossen (1979) (henceforth Heine et al. 1979) with additional proposals from Ehret (1970, 1971, 1974) that there once was an East

1 The data that the article is based on are from published sources referred to in the references to the discussion in the Appendix. The research was sponsored by a grant (406.18.TW.013) from the Dutch Research Council www.lheaf.org; there is no conflict of interests. We are grateful to the very useful comments from the anonymous reviewers and inspiring discussions with Maarten Kossmann and members of the LHEAf-team. We thank Nina van der Vlugt for extensive editorial assistance.

DOI 10.15460/auue.2025.98.1.368

Cushitic language from which words were borrowed into Proto South Nilotic. The existence of this putative language, called Proto Baz by Heine et al. (1979), can only be inferred from these loans. Heine et al. (1979) identify 37 Proto Baz loanwords in South Nilotic, which amounts to almost one-fourth of the reconstructed Proto South Nilotic lexicon.

The Proto Baz hypothesis has received considerable attention in the literature (39 citations up to 2023 according to Google Scholar), mostly by linguists, but also by archaeologists and historians. This is not surprising given the scarcity of data on the prehistory of the region. Less expected may be that Proto Baz is not mentioned in overviews of Cushitic such as Sasse (1981), Tosco (2000), and Appleyard (2012), nor in the detailed overview of Afroasiatic by Hayward (2000) or the overview of East Cushitic by Tosco (2020); an overview of Cushitic that does mention it is Mous (2012). In spite of this, the Proto Baz hypothesis has never been discussed in detail, and the present article seeks to address this lacuna. Further, we discuss loans from Oromoid and Yaaku-Dullay that have been mentioned in the endnotes of Heine et al. (1979) and in Ehret (1970, 1971).

Heine et al. (1979) view Proto Baz as a separate branch of Omo-Tana within East Cushitic (Fig. 1). The language would have been part of a third major southward movement of Omo-Tana-speaking peoples from the Highlands of Ethiopia into the plains of northern Kenya, in addition to the movements linked to present-day East and West Omo-Tana (Heine et al. 1979: 81–82). Proto Baz speakers would have settled on the north-western side of Lake Turkana (hence the name of their language, from Proto Omo-Tana *baz ‘lake’), where they would have been in contact with Proto South Nilotic.

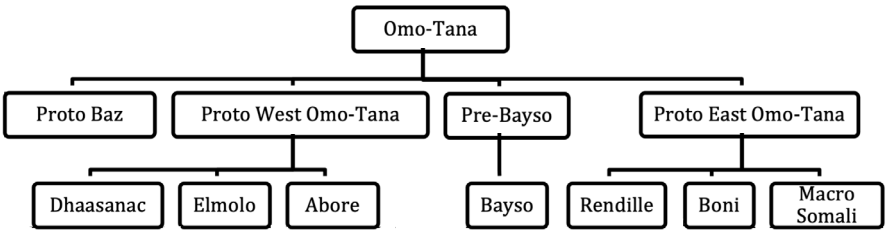


Figure 1: Omo-Tana subclassification (adapted from Heine et al. 1979: 80)

In order to propose a single putative language as the source of loans, it is necessary to identify a set of features that uniquely characterise

DOI 10.15460/auue.2025.98.1.368

that language, possibly over some period of time. For a number of the proposed loan words this is problematic. Further, to postulate a separate language as the source of loans, the source items must be different from the known languages and their earlier forms. This is not argued for in detail in Ehret (1970, 1971, 1974) and Heine et al. (1979). The latter, however, does discuss the adaptations that Proto Southern Nilotic made to the loans in order to fit in its sound system, and it is this article that is at the centre of our discussion.

We discuss the two points:

(i) Do all the putative loans from Proto Baz share the same origin?

(ii) Are there sound changes that warrant a separate language, different from other known or reconstructed Cushitic languages?

Our conclusion is that both points have to be answered in the negative: All loans do not come from the same Cushitic language, and sound changes do not warrant a separate Cushitic language (section 2). Rather, we identify several Cushitic donor languages for the so-called Proto Baz loans in South Nilotic (section 3), and these loans are likely to have entered South Nilotic at different times (sections 5, 6, 7). The full discussion of each of the lexical transfers are presented in the Appendix which contains all lexical items proposed by Heine et al. (1979) including the ones from Ehret (1970, 1971) that they refer to in their endnotes. The items are alphabetically ordered and start out with the original proposal. We label these putative cognate sets with an approximation of borrowed item presented in CAPITALS. Language specimen are italicised but reconstructed items are not.

2 A separate language Proto Baz?

Heine et al. (1979) suggest the following loan adaptations of the Cushitic loan words:

(i) All obstruents are interpreted as voiceless, in line with the absence of voice opposition in obstruents in Proto South Nilotic and Proto Kalenjin (Rottland 1982: 216, 231)

(ii) The pharyngeals are deleted or replaced as they do not exist in Proto South Nilotic

(iii) Long consonants are shortened in line with the absence of geminates in Proto South Nilotic

DOI 10.15460/auue.2025.98.1.368

These adaptations are logical and straightforward. However, we consider the replacement of the voiced pharyngeal by *s* as phonetically too distinct and highly unlikely and we do not accept the three cognate sets that display it. The most common replacement for the voiced pharyngeal in Cushitic languages is the glottal stop (cf. Sasse 1979: 35).[2] The loss of the voiceless pharyngeal *ħ* is attested only word-finally in the loans, and we consider this plausible in that environment.

The sound inventory of Proto Baz is arrived at by taking into account these adaptations and is largely identical to Proto East Omo-Tana ('Proto Sam' in Heine et al. 1979). Differences lie in the following two rules from the perspective of Proto East Omo-Tana (for Proto East Omo-Tana, cf. Heine 1978, 1981; Lamberti 1986; for Proto South Nilotic, Rottland 1982):

1. Raising of Proto East Omo-Tana *a > Proto Baz *ɛ/e in the context of pharyngeal consonants.
2. Split of Proto East Omo-Tana *r into two rhotics, Proto Baz *r and *R.

We do not accept the first change as one that warrants setting up a separate putative language, since all Cushitic languages that have these pharyngeals show allophonic raising of vowels in the context of pharyngeals. Heine et al. (1979) posit a separate phoneme /ɛ/ for Proto Baz (1979: 76, 77, 80), but this is not warranted as it occurs only in pharyngeal environments, to the exclusion of *a* and *e*. Moreover, no known Cushitic language has a six-vowel system of the type reconstructed for Proto Baz (*i*, *e*, *a*, *o*, *u*, plus *ɛ*). Rather, borrowing into South Nilotic simply seems to have followed the Cushitic phonetics and not the phonemic abstraction in the source language. Proto East Omo-Tana and Proto Omo-Tana, Proto Oromoid, Proto Yaaku-Dullay all must have had a voiced pharyngeal; only Proto West Omo-Tana has the glottal stop for the original voiced pharyngeal.

The Proto East Omo-Tana *r to Proto Baz *R correspondence is based on two cognate sets, *maR 'calf' and *aR 'male bovine'. The first one is possibly not part of the Proto Baz borrowing event because, as Heine et al. (1979) remark in their endnote 41, the root is attested

2 Notably, Heine et al. (1979) do not propose a glottal stop in the phoneme inventory of Proto Baz, but most likely this is an omission resulting from the orthographic convention not to write a word-initial glottal stop even though it is phonemic in Cushitic languages.

DOI 10.15460/auue.2025.98.1.368

wider in East Sudanic and is likely inherited from Eastern Sudanic in Proto Southern Nilotic. The root for 'male bovine', *aR, is linked to the Western Omo-Tana languages Arbore and Dhaasanac. This suggests an alternative scenario of a different borrowing event, not from Proto Baz. Hence the suggested split cannot be taken to differentiate Proto Baz from other Omo-Tana languages in our view. The correspondence Proto East Omo-Tana *r* to Proto Baz *r* is valid in our view but on evidence that differs from the proposed evidence in Heine et al. (1979). They based this correspondence on three cognate sets, AFARTAM 'forty', REEREE 'bat', and GORO 'fur'. The last one we find unconvincing (see Appendix); the set for 'bat' lacks substantiation for Proto East Omo-Tana, and the set for 'forty' in Heine et al. (1979) is neither substantiated by the reconstructions of South Nilotic (Rottland 1982) nor of Proto East Omo-Tana by Heine (1978) and Lamberti (1986); we did find additional evidence to suggest South Nilotic borrowing this item from proto Omo Tana, see Appendix. However, the correspondence of Proto East Omo-Tana *r and Proto Baz *r is likely based on other evidence, as we argue in the Appendix for items such as DERI 'clay pot', BURI 'red', INSIR 'louse' and ROOB 'rain'. Concluding, the unproblematic correspondence *r*=*r* is irrelevant for recognising a separate donor language. The existence of an extra consonant *R* in Proto Baz is not needed if we allow transfer from West Omo-Tana into Proto Kalenjin in the instance of *aR* 'male bovine'.

3 A differentiation of sources

Heine et al. (1979) give both Proto South Nilotic and Proto Kalenjin as evidence for early South Nilotic contact with East Cushitic. They pose a now extinct East Cushitic language Proto Baz on the assumption of transfer into Proto South Nilotic. The essence of posing such a putative language is the hypothesis that the set of words that was transferred from Cushitic into South Nilotic belonged to one and the same language. Once we let the assumption go and allow for several donor sources, a more complex picture emerges that stays closer to the available evidence. In our discussion of the cognate evidence, we will entertain the less speculative hypothesis, namely that loans came from existing Cushitic languages including reconstructed stages, and that there is no need for setting up a separate language.

DOI 10.15460/auue.2025.98.1.368

The available evidence for each item is discussed in the Appendix and the results summarised below in section 3.2. Before that, it is crucial to single out recurring challenges in the historical interpretation of the evidence.

3.1 Challenges for interpretation of the evidence

There are several recurring challenges in the historical interpretation of Cushitic loans in South Nilotic: formal ambiguity (i), direction of transfer (ii), conflicting reconstructions (iii), target language in South Nilotic (iv), formal distance between Proto Baz and South Nilotic (v), missing evidence in Cushitic (vi), and missing discussion (vii). These are discussed below.

(i) Formal ambiguity

For some of the lexemes, it is difficult to decide what the Cushitic source must have been. For ROOB 'rain', for example, this is due to its wide distribution in Cushitic with little variation in shape; the root QUUT 'scratch' could be equally well West Omo-Tana as East Omo-Tana, or even Oromoid; and TOR 'spear' may be from West Omo-Tana or from (Pre-)Yaaku.

(ii) Direction of transfer

A number of items are restricted in their distribution in Cushitic, which raises the question of the direction of borrowing. This is not addressed in Heine et al. (1979). It is conceivable that the word for 'red' BURI was transferred from South Nilotic into Proto West Omo-Tana as it is limited to West Omo-Tana within Cushitic. However, exactly where in the evolution of West Omo-Tana the borrowing would have occurred is hard to determine.

(iii) Conflicting reconstructions

A number of the proposals in Heine et al. (1979) are difficult to evaluate because a Proto South Nilotic, Proto Kalenjin or Common Datooga form is given for which we find no support in daughter languages, and these proposals eventually did not make it to Rottland (1982, 1989, 1997).[3] This includes the three items AFARTAM 'forty', ILAAL 'to look at', KOR 'bell', and MAQAL 'young goat/sheep'. We evaluate whatever additional evidence we could find for these items in the Appendix.

3 Rottland (1989) and (1997) contain minor extensions and corrections of Rottland (1982), but without presenting the evidence on which they are based. For this reason, Rottland (1982) is our main source for South Nilotic.

DOI 10.15460/auue.2025.98.1.368

(iv) Target language in South Nilotic

The item NUG 'to suck' is only linked to Datooga in Heine et al. (1979). There is no evidence (yet) for such a root in Kalenjin, which leaves two options: transfer from Cushitic into Datooga, or into Proto South Nilotic and subsequent loss in Kalenjin. Where Heine et al. (1979) seem to go for the second option, the later work by Rottland (1982) apparently opts for the first.

(v) Formal distance between Proto Baz and South Nilotic

We find a number of the Proto Baz proposals unconvincing on formal grounds. For example, the proposal for 'grass' requires the unlikely correspondence of voiced pharyngeal fricative to *s*, as in Proto Kalenjin *suus- from an alleged Cushitic source related to Proto East Cushitic *ʕawʃ-/*ʕayʃ- 'grass' (Sasse 1979). The proposal for GORO 'feathers, fur' requires explanation of the extra final consonant in Nilotic but is otherwise promising. Details on our evaluation of these transfers are to be found in the Appendix.

(vi) Missing evidence in Cushitic

We still have to find evidence on the East Omo-Tana side for some items that are mentioned in Heine et al. (1979) but that are not present in Heine (1978) and Lamberti (1986): ZIG 'mud, dung', SAANI 'lover'. We report on our unsuccessful search in the Appendix.

(vii) Missing discussion

One word mentioned in the final table in Heine et al. (1979), IRI 'smoke', is actually not discussed by them in their article and we failed to find cognates with Cushitic, as we report in the Appendix. Another word, WAAR 'kid', is mentioned as a tentative candidate in their endnote 41, and we discuss it in the Appendix.

3.2 Cushitic donor languages

As mentioned above, evidence suggests that the alleged Proto Baz loans in South Nilotic stem from a range of Cushitic languages rather than a single one, Proto Baz. The likely source languages include Proto Omo-Tana, Proto East Omo-Tana, Proto West Omo-Tana, Proto Yaaku-Dullay and Proto Oromoid, and their daughters. It will be noted that Proto Baz is not among them, as we do not find compelling evidence that it ever existed. At the receiving end, we distinguish between Proto Kalenjin and Proto South Nilotic and their daughters. In some instances, the Cushitic loan is apparent in Common Datooga. In those cases, we discuss the two logical options: transfer into Proto

DOI 10.15460/auue.2025.98.1.368

Datooga and transfer into Proto South Nilotic but subsequent loss in Proto Kalenjin. The former option has the challenge of the geographical distance, in as far as the homeland of Proto Datooga can be established.

4 Loans from Proto Omo-Tana

In this section we discuss the loans that could have Proto Omo-Tana as source. In the next section we discuss those that are specifically from Proto East Omo-Tana. However, many of the items in this section on Proto Omo-Tana could also be postulated to have East Omo-Tana as source. On the Nilotic side, the transfers in this section are to Proto South Nilotic, or possibly to that level, while those in the next section are to a lower level, Proto Kalenjin. In this section, we also include the Cushitic-Nilotic transfers that may have happened at even earlier periods into Proto Nilotic.

Loans from Proto Omo-Tana predate the developments that characterise West Omo-Tana and East Omo-Tana respectively. In almost all cases the transfers were into Proto South Nilotic. In rare cases the item reconstructs only to Proto Kalenjin in South Nilotic; for reasons of patterning, it can be surmised that a subsequent innovation happened in Datooga (cf. the uneven decimals, such as SOZZOM in the Appendix).

This set of Proto Omo-Tana loans is further conspicuous for consisting almost exclusively of numerals (six to ten, and most decimals, 100). It is remarkable that South Nilotic basic numbers from six to ten are Cushitic in origin. It is reasonable to assume that this constitutes transfer from a single source. The transfer of these numbers differs from the transfer from East Omo-Tana into Proto Kalenjin. One reason is that the Cushitic numbers are not restricted to Kalenjin but all are also present in Datooga. Yet Rottland (1982) does not reconstruct them to Proto South Nilotic level; presumably irregularities prevent him from doing so.

Another aspect of the transfer of these numerals to consider is the source. The source does not point uniquely to East Omo-Tana but could just as well be at the higher Omo-Tana level and if we assume it is East Omo-Tana, we need to assume that it predates the inner East Omo-Tana sound changes that affect *z (‘seven’ and ‘eight’) and *m. The shapes of the loans in South Nilotic must have predated the

DOI 10.15460/auue.2025.98.1.368

z > *d* sound change of Northern Somali and *z* > *y* in Rendille and Southern Somali which points to an early loan, either at Proto East Omo-Tana level or earlier.

In addition to these basic numbers, the expressions for '30', '40', and '50', '60', '80' are from a Cushitic source. Somali has single expressions for these numbers, some of which can be shown to derive from a construction numeral + 'ten' but amalgamated and reduced to one unit, as we also see in Agaw, Sidamo (Highland East Cushitic), Afar, Bayso (Zelealem 2003). These are retentions in Somali from (East) Omo-Tana while Rendille has innovated a syntactic formation using an expression in which the unit for 'ten' is followed by the word for 'three', 'four', etc.

Blažek (2018: 53) shows that these basic numbers, two, three, five to ten were also transferred into Southeast Surmic and proposes that Southeast Surmic is the source for the presence of these ultimately Cushitic roots in East Nilotic Maa, arguing on the basis of phonetic similarity and linguistic geography. He argues that the Cushitic source for this transfer into Southeast Surmic is (late) West Omo-Tana.

Other early transfers from Cushitic to Nilotic are in our view QUUT 'scratch' which has cognates in East Nilotic, ROOB 'rain', which can be from any level of East Cushitic into Proto South Nilotic, MUR 'cut, circumcise, marry', which is reconstructable to the Proto Nilotic level, AM 'eat' from Proto (East) Omo-Tana to Proto South Nilotic. The item PATAI 'back' is likely to be a transfer in the other direction, from Proto South Nilotic to Proto Omo-Tana.

The transfer of this large set of numerals from Proto Omo-Tana to Proto South Nilotic begs for historical cultural interpretations. It is clear the Cushitic language was dominant in the area as the transfers are from Cushitic to Nilotic. Borrowing (higher) numbers is a common phenomenon cross-linguistically but nevertheless it is hard to guess why the Nilotic speakers would use the Cushitic numerals: Assuming that there was no money economy, is the necessity for higher numbers related to larger herds of domestic animals, or to the Cushitic age-set system? Many of the non-numeral transfers are verbs: ROOB 'to rain' is borrowed as a verb in Proto South Nilotic, AM 'eat', MUR 'cut' QUUT 'scratch' are all verbs and quite basic verbs hence replacive rather than additive borrowings; these are facts that suggest intense contact, as verbs are less prone to borrowing.

DOI 10.15460/auue.2025.98.1.368

5 Loans from East Omo-Tana

In our re-assessment of the proposed loans in Proto Baz, we have identified only three that are likely to be transfers from Proto East Omo-Tana to South Nilotic, and actually specifically to Proto Kalenjin: INSIR ‘louse’, ɗERI ‘claypot’, and SUBEEN ‘ewe’ (see Appendix). For KAS ‘to understand’ we propose transfer in the other direction, from Proto Kalenjin to Proto East Omo-Tana. If the transfers are restricted to these, not much can be concluded other than there was contact between Proto East Omo-Tana and Proto Kalenjin involving two cultural items, ‘clay pot’ and ‘ewe’. However, keep in mind that some of the transfers that were included in the preceding section, could also be considered to belong here.

The transfer of MAQAL ‘young goat/sheep’ was in our view later and at a lower level, from Rendille into Pokot. In the case of SAANI ‘lover’ we propose the other direction, also later and at a lower level, from a yet to be specified Kalenjin language into Somali.

6 Loans from West-Omo-Tana

There are a number of items that seem to be transferred from West Omo-Tana (previously known as Galaboid) containing Dhaasanac, Arbore and Elmolo. For these items the Proto West Omo-Tana form is closer to the South Nilotic form compared to the East Omo-Tana form.

Map 1 in Heine et al. (1979) posits Proto West-Omo-Tana closer to Proto South Nilotic and Proto East Omo-Tana further to the east. Tosco (2015: 105) argues why Elmolo is an unlikely source for transfer into Nilotic: “While the presence of Cushitic words in Southern (and Eastern) Nilotic is a fact, there is scarcely any specific evidence (other than geographic proximity) linking these loans to any fishing community in the Lake Turkana area and to the Elmolo in particular.” Interestingly, this contact with Proto West Omo-Tana may have been earlier than the contact of Proto East Omo-Tana with Proto Kalenjin because the Proto West Omo-Tana items can be proposed to have entered Proto South Nilotic rather than Proto Kalenjin. This is the case for AR ‘male bovine’, QUAR ‘male goat’, METEH ‘head’, and possibly TOR ‘spear’. MALAB ‘honey’ is only attested in Datooga; however, in order to suggest a plausible contact scenario

DOI 10.15460/auue.2025.98.1.368

in terms of time depth and geographical proximity, we suggest an additional lexical replacive borrowing for 'honey' in Proto Kalenjin to account for the absence of this root in Kalenjin, see SEKEM in the Appendix. The root MAR 'calf' is likely to be a transfer from Nilotic or Surmic into Proto West Omo-Tana.

Although these items are few in number, they require that Proto West Omo-Tana and Proto South Nilotic were contemporary and in geographical proximity. The proposed transfers with two of them referring to male domestic animals suggests that certain aspects of animal husbandry (possibly breeding) were innovated among the South Nilotes from contact with Proto West Omo-Tana. The third concept 'head' must have been replacive rather than additive, suggesting more than superficial contact. The cultural dominance of Proto West Omo-Tana is also evident from their influence on Southeast Surmic. Blažek (2018: 53, 41) shows that Southeast Surmic borrowed the numerals 2–10 (except 4) from East Cushitic and he compares it in particular to Proto West Omo-Tana.

7 Loans from Yaaku-Dullay

Two to four items are linked to Yaaku-Dullay, a branch of East Cushitic also known as Peripheral-East-Cushitic (Tosco 2020) and referred to as Yaaku-Qawko in Heine et al. (1979) following Hayward (1978). One item, SEKEM 'bee/honey' is found in Proto Kalenjin, the other, EU 'night', in Datooga. The data is too limited to allow for anything more than speculation as to whether they entered at the Proto South Nilotic level and subsequently got lost in one of the primary branches of South Nilotic, or whether Proto Kalenjin and Proto-Datooga were the receiving languages. Possibly TOR 'spear' belongs here too, although it could also be a transfer from Proto West Omo-Tana. And BOD 'get lost' appears to be a loan into Pokot from Dullay.

The history of Yaaku-Dullay and its speakers is still little known but discussed in Sasse (1979: 54), Tosco & Sands (2022), Hayward (1978), and Ehret (2008). This is one of the areas where progress could significantly alter our understanding of the history of the region. South Nilotic, however, does not seem to have had intensive contact with this Cushitic subgroup.

DOI 10.15460/auue.2025.98.1.368

8 Loans from Oromo(id)

Heine et al. (1997: 88) report that Ehret suggested to them some correspondences between Southern Nilotic and Oromoid and that he proposed transfer from Oromoid through Proto Baz to the Southern Nilotes. Four etymologies are mentioned, and we discuss them in the Appendix. We confirm RAARO 'leather container', KOLKOC 'tortoise' as likely transfers from an Oromoid source to Proto Kalenjin. For TWAAL 'cowbell' the situation is unclear. Finally, we are not convinced that the roots for 'black ant' under RIRIM are actually cognate.

The transfers seem to be into Proto Kalenjin; we found no attestations in other branches of South Nilotic for these items. This suggests that the Oromoid source was roughly contemporary with Proto Kalenjin, and, if Proto East Omo-Tana was indeed in contact with Proto Kalenjin, also contemporary with Proto East Omo-Tana. It is difficult however to establish the level of the "Oromoid" source. For Oromo RAARO we have no other cognates within Oromoid, and for KOLKOC various levels within Oromoid are possible as source.

9 Summary and conclusion

Our revision of the evidence presented in Heine et al. (1979) in the light of new data and insights results in the following picture of Cushitic - Nilotic contacts in Northern Kenya.

There was transfer from Proto **Omo-Tana** into Proto South Nilotic, showing cultural dominance of Proto Omo-Tana in particular because of the paradigm of decimal numbers that Proto South Nilotic borrowed from Proto Omo-Tana but also other numbers, 6–10. In addition, at least four relatively basic verbs in Proto South Nilotic are of Proto Omo-Tana provenance and both aspects, verb and basicness, point to intense contact.

There was some transfer from Proto **West Omo-Tana** into Proto South Nilotic. Although the number of borrowed words is low, the fact that two of them refer to male domestic animals may be a significant indication for the nature of the West Omo-Tana cultural influence on Proto South Nilotic.

There was some transfer from Proto **East Omo-Tana** into Proto Kalenjin. This is partly a consequence of our decision to take those items that could have been transferred at earlier stages at donor

DOI 10.15460/auue.2025.98.1.368

and receiving side to be evidence of that earlier contact event with the advantage of regularising Omo-Tana – South Nilotic versus East Omo-Tana – Kalenjin contact events.

There was very little transfer from an **Oromoid** source into Proto Kalenjin; too little to draw any conclusions from as yet and the same is true for scarcity of evidence for **Yaaku/Dullay** transfer into any of the Southern Nilotic languages.

We do not accept the evidence for a separate Omo-Tana language Baz on the basis of the presented Cushitic loans in South Nilotic. Not assuming one source for these loans made it possible for us to suggest a number of different contact scenarios. These scenarios need further substantiation and would ideally be linked to results from other disciplines. We present Heine et al.'s (1979) map 1 here as Figure 2. We refrain from presenting a map situating the proto-languages and their contact zones because that would already involve bringing in non-linguistic considerations. Instead, we offer a schematic "map", Figure 3, representing our main conclusions on language contact in earlier times between Cushitic and South Nilotic.

We envisage that this revision of earlier research on Cushitic – Nilotic contacts marks the beginning of re-appraisal of such work. Many more lexical similarities between the two families need to be uncovered and analysed. We are currently working on the intense South Cushitic South Nilotic contact. We hope that this work inspires research on contact beyond the lexicon and thinking on the socio-historical scenarios for these contact events.

DOI 10.15460/auue.2025.98.1.368

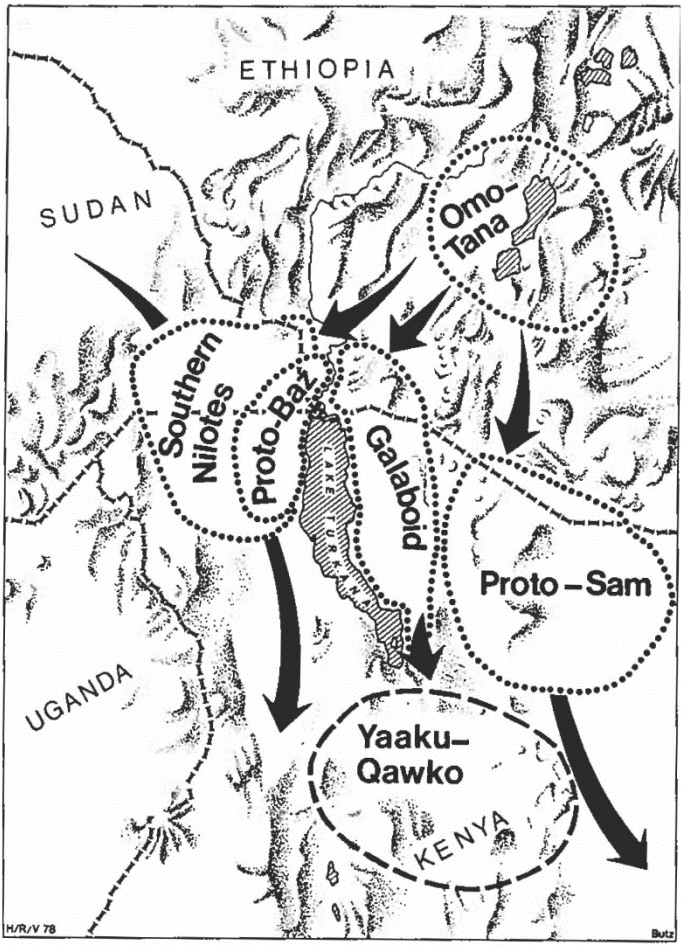


Figure 2: Nilotic-Cushitic contacts between 1000 B.C. and 100 A.D. (Heine et al. 1979: 84).

DOI 10.15460/auue.2025.98.1.368

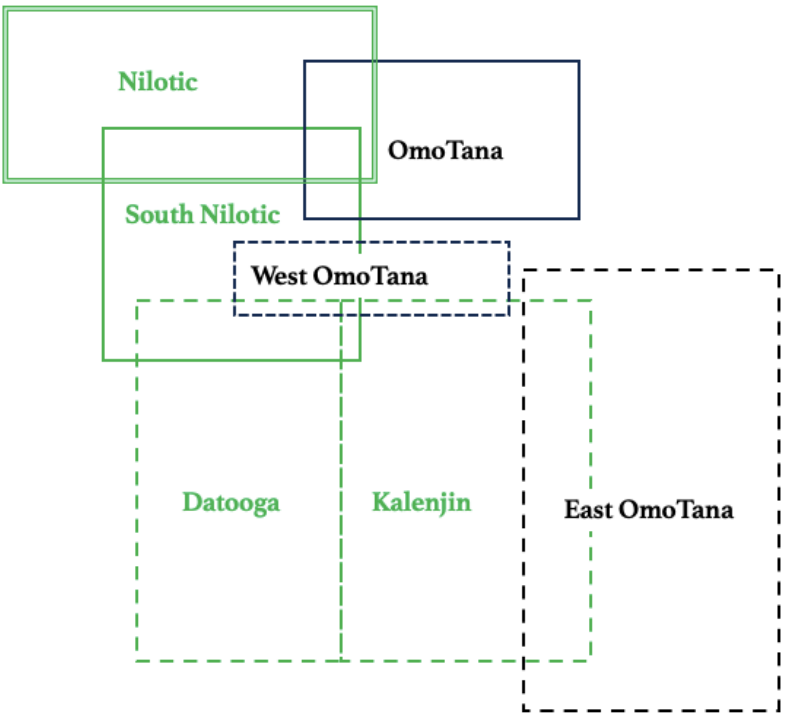


Figure 3: Schematic map of the Cushitic and South Nilotic contacts based on the reassessed lexical evidence

References

Appleyard, David L. 2006. *A comparative dictionary of the Agaw languages*. (Kuschitische Sprachstudien 24). Köln: Rüdiger Köppe Verlag.

Appleyard, David L. 2012. Cushitic. In Lutz Edzard (ed.), *Semitic and Afro-asiatic: Challenges and opportunities*. Wiesbaden: Otto Harrassowitz. 199–292.

Baader, Christian. nd. Yakunte-English dictionary. Unpublished manuscript.

Banti, Giorgi & Moreno Vergari. 2005. A sketch of Saho grammar. *Journal of Eritrean Studies* 4(1–2). 100–131.

Barrett, Anthony J. 1988. *English-Turkana dictionary*. Nairobi: Macmillan Kenya (Publishers) Limited.

Beech, Merwyn Worcester Howard. 1911. *The Suk: their language and folklore*. Oxford: Clarendon Press.

DOI 10.15460/auue.2025.98.1.368

Black, Paul David. 1973. Preliminary draft of a Gidole dictionary. Unpublished manuscript.

Black, Paul David. 1974. *Lowland East Cushitic: Subgrouping and reconstruction*. New Haven: Yale University.

Black, Paul David. 1976. Werizoid. In Marvin Lionel Bender (ed.), *The Non-Semitic languages of Ethiopia*. Michigan: Michigan University.

Black, Paul David & Shako Otto. 1973. *First draft of a Konso dictionary*. Murray Hill, New Jersey: Bell Labs.

Blažek, Václav. 2018. Cushitic numerals. *Folia Orientalia* 55. 33–60. https://doi.org/10.24425/for.2018.124678.

BLR3: Bastin, Yvonne, André Coupez, Evariste Mumba & Thilo C. Schadeberg (eds.). 2002. Bantu lexical reconstructions 3 / Reconstructions lexicales bantoues 3. http://linguistics.africamuseum.be/BLR3.html

Conti Rossini, Carlo. 1904. Appunti sulla lingua khamta dell'Averghellé. *Giornale della Societa Asiatica Italiana* 17. 183–242.

Crazzolara, Pasquale. 1978. *A study of the Pokot (Suk) language: Grammar and vocabulary*. Bologna: Editrice Missionaria Italiana.

Creider, Jane Tapsubei & Chet A. Creider. 2001. *A dictionary of the Nandi language.* (Nilo-Saharan: Linguistic Analyses and Documentation 16). Köln: Rüdiger Köppe Verlag.

Dhoorre, Cabdulqaadir Salaad & Mauro Tosco 1998. Somali ideophones. *Journal of African Cultural Studies* 11(2). 125–156

Dimmendaal, Gerrit J. 1984. Review of: Rainer Vossen, The Eastern Nilotes: Linguistic and historical reconstructions. *Journal of African Languages and Linguistics* 6. 67–75.

Dimmendaal, Gerrit J. 1988. The lexical reconstruction of Proto-Nilotic: A first reconaissance. *Afrikanistische Arbeitspapiere* 16. 5–67.

Ehret, Christopher. 1970 [1969]. *The Southern Nilotes to 1600 A.D.: A linguistic approach to East African history*. PhD Northwestern University, 1969. Ann Arbor: University Microfilms Inc.

Ehret, Christopher. 1971. *Southern Nilotic history: Linguistic approaches to the study of the past.* Evanston: Northwestern University Press.

Ehret, Christopher. 1974. *Ethiopians and East Africans: The problem of contacts.* (Nairobi Historical Studies 3). Nairobi: East African Publication House.

Ehret, Christopher. 1980. *The historical reconstruction of Southern Cushitic phonology and vocabulary.* (Kölner Beiträge zur Afrikanistik 5). Berlin: Dietrich Reimer.

DOI 10.15460/auue.2025.98.1.368

Ehret, Christopher. 1981. Revising Proto-Kuliak. *Afrika und Übersee* 64. 81–100.

Ehret, Christopher. 2008. Yaakuan and Eastern Cushitic: A historical linguistic overview. In Gábor Takács (ed.), *Semito-Hamitic Festschrift for A. B. Dolgopolsky and H. Jungraithmayr*. Berlin: Dietrich Reimer. 128–141.

Ehret, Christopher & Derek Nurse 1981. The Taita Cushitic. *Sprache und Geschichte in Afrika* 3. 125–168.

Gragg, Gene B. 1982. *Oromo dictionary* (Monographs from the Committee on Northeast African Studies, Ethiopian Series 12). East Lansing: African Studies Center, Michigan State University.

Griscom, Richard (ed.). 2018. *Ng'asheeg Ásímjéég: Kamusi fupi ya lugha ya Kiasimjeeg / Short dictionary of the Asimjeeg Datooga language*. With contributions by Michael Bacheng, Yonah Ndege, Yohani G. Mangi, Gidajur Gitonyod, Bwana Karatu, Berta Duru, Eliya Bacheng, Agustino Mangi. Available at datooga-griscom-0378 | Endangered Languages Archive (elararchive.org).

Griscom, Richard. 2019. *Topics in Asimjeeg Datooga verbal morphosyntax*. Oregon: University of Oregon Dissertation.

Hayward, Richard J. 1978. The Qawko dialects and Yaaku. *ABBAY: Documents pour servir à la histoire de la civilisation Éthiopienne* 9. 59–70.

Hayward, Richard J. 1984. *The Abore language: A first investigation including a vocabulary*. (Kuschitische Sprachstudien 2). Hamburg: Helmut Buske.

Hayward, Richard J. 2000. Afroasiatic. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*. Cambridge: Cambridge University Press. 74–98.

Heine, Bernd. 1978. The Sam languages: A history of Rendille, Boni and Somali. *Afroasiatic Linguistics* 6(2). 1–93.

Heine, Bernd. 1980. *Elmolo*. (The Non-Bantu languages of Kenya). Berlin: Dietrich Reimer.

Heine, Bernd. 1981. Some cultural evidence on the early Sam-speaking people of Eastern Africa. *Sprache und Geschichte in Afrika* 3. 169–200.

Heine, Bernd, Franz Rottland & Rainer Vossen. 1979. Proto-Baz: Some aspects of early Nilotic-Cushitic contacts. *Sprache und Geschichte in Afrika* 1. 75–91.

Hudson, Grover. 1989. *Highland East Cushitic dictionary*. (Kuschitische Sprachstudien 7). Hamburg: Helmut Buske.

DOI 10.15460/auue.2025.98.1.368

Kießling, Roland & Maarten Mous. 2003. *The lexical reconstruction of West-Rift Southern Cushitic.* (Kuschitische Sprachstudien 21). Köln: Rüdiger Köppe Verlag.

Lamberti, Marcello. 1986. *Die Somali-Dialekte: eine vergleichende Untersuchung.* (Kuschitische Sprachstudien 5). Köln: Rüdiger Köppe Verlag.

Lamberti, Marcello. 1988. *Kuliak and Cushitic: A comparative study.* (Studia Linguarum Africae Orientalis 3). Heidelberg: Carl Winter.

Leslau, Wolf. 1991. *Comparative dictionary of Ge'ez (Classical Ethiopic): Ge'ez-English/English-Ge'ez with an index of the Semitic roots*. Wiesbaden: Otto Harrassowitz.

Mous, Maarten. 2012. Cushitic. In Zygmunt Frajzyngier & Erin Shay (eds.), *The Afroasiatic languages*. Amsterdam & Cambridge: Cambridge University Press. 342–422.

Ohta, Itaru. 1989. A classified vocabulary of the Turkana in northwestern Kenya. *African Study Monographs, Supplementary Issue* 10. 1–104.

Omondi, Odero Erick & Odero Isack Otieno. 2008. Elmolo-English dictionary. With Lengosira Andrea, Lengutuk Asunta and Akolong Joseph. Unpublished MS, ms.

Orkaydo, Ongaye Oda. 2013. *A grammar of Konso.* (PhD dissertation, Leiden University). Utrecht: LOT dissertation series.

Payne, Doris J. & Leonard Kotikash. 2005, 2008. Maa (Maasai) Dictionary. https://darkwing.uoregon.edu/~dlpayne/Maa%20Lexicon/lexicon/main.htm.

Pillinger, Steve & Letiwa Galboran. 1999. *A Rendille dictionary: including a grammatical outline and an English-Rendille index.* (Kuschitische Sprachstudien 14). Köln: Rüdiger Köppe Verlag.

Rottland, Franz. 1982. *Die Südnilotischen Sprachen.* (Kölner Beiträge zur Afrikanistik 7). Berlin: Dietrich Reimer.

Rottland, Franz. 1989. Southern Nilotic reconstructions. In Marvin Lionel Bender (ed.), *Southern Nilotic reconstructions.* (Topics in Nilo-Saharan Linguistics). Hamburg: Helmut Buske. 219–231.

Rottland, Franz. 1997. Lexical reconstruction within Nilotic. *Afrikanistische Arbeitspapiere* 52. 139–173.

Sacleux, Charles. 1939. *Dictionnaire swahili-français.* (Travaux et mémoires de l'Institut d'Ethnologie 36–37). Paris: Institut d'Ethnologie.

Said 2013. *Afmaal Somali-English Dictionary*. www.afmaal.com

Samburu-English Online Dictionary. Webonary (SIL International), 2019. https://www.webonary.org/samburu/.

DOI 10.15460/auue.2025.98.1.368

Sasse, Hans-Jürgen. 1979. The consonant phonemes of Proto-East-Cushitic (PEC): A first approximation. *Afroasiatic Linguistics* 7(1). 1–67.

Sasse, Hans-Jürgen. 1981. Die kuschitischen Sprachen. In Bernd Heine, Thilo C. Schadeberg & Ekkehard Wolff (eds.), *Die Sprachen Afrikas*. Hamburg: Helmut Buske. 187–216.

Sasse, Hans-Jürgen. 1982. *An etymological dictionary of Burji.* (Kuschitische Sprachstudien 1). Hamburg: Helmut Buske.

Schrok, Terrill. 2017. *The Ik language: Dictionary and grammar sketch.* (African Language Grammars and Dictionaries 1). Berlin: Language Science Press. 10.5281/zenodo.344792.

Sosal, Ahmed. In prep. *PhD thesis (in preparation)*. Leiden: Leiden University.

Stroomer, Harry J. 2001. *A concise vocabulary of Orma Oromo (Kenya): Orma-English, English-Orma.* (Kuschitische Sprachstudien 17). Köln: Rüdiger Köppe Verlag.

Sweetser, Eve. 1990. *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge. Cambridge University Press.

Tosco, Mauro. 1994. The historical syntax of East Cushitic: A first sketch. In Thomas Bearth, Wilhelm J. G. Möhlig, Beat Sottas & Edgard Suter (eds.), *Perspektiven afrikanistische Forschung. Beiträge zur Linguistik, Ethnologie, Geschichte, Philosophie und Literatur / X. Afrikanistentag (Zürich, 23–25 September 1993)*. Köln: Rüdiger Köppe Verlag. 415–440.

Tosco, Mauro. 2000. Cushitic overview. *Journal of Ethiopian Studies* 33(2). 87–121.

Tosco, Mauro. 2001. *The Dhaasanac language* (Kuschitische Sprachstudien 17). Köln: Rüdiger Köppe Verlag.

Tosco, Mauro. 2015. From Elmolo to Gura Pau: a remembered Cushitic language of Lake Turkana and its possible revitalization. *Studies in African Linguistics* 44(2). 101–135.

Tosco, Mauro. 2020. East Cushitic. In Rainer Vossen & Gerrit J. Dimmendaal (eds.), *The Oxford handbook of African languages*. London: Oxford University Press. 290–299.

Tosco, Mauro & Bonny Sands. 2022. Early East African and peripheral East Cushitic: Foragers and pastoralists in early East Africa. Presented at the Rift Valley Network Webinar Series. https://doi.org/10.5281/zenodo.7440446.

DOI 10.15460/auue.2025.98.1.368

Vergari, Moreno & Roberta Vergari. 2007. *A basic Saho-English-Italian dictionary. Revised version of the original 2003 edition.* Asmara: Sabur Printing Services.

Vossen, Rainer. 1982. *The Eastern Nilotes: Linguistic and historical reconstructions.* (Kölner Beiträge zur Afrikanistik 9). Berlin: Dietrich Reimer.

Wedekind, Klaus, Feki Mahamed, Mohammed Talib, Abuzeinab Musa, Ibrahim Mohammad, Oshe Tahir & Andy. 2021. *Beja dictionary (Beja-English-Arabic / English-Beja-Arabic)*. (Ed.) Charlotte Wedekind & Jonah Wedekind. Köln: Rüdiger Köppe Verlag.

Yibeltal Getahun. 2018. *Documentation of Mositacha and description of its morphology*. Addis Ababa: Addis Ababa University PhD thesis.

Yri, Kjell Magne & Steve Pepper. 2019. *Sidaama dictionary.* (Dictionaria 6). 10.5281/zenodo.5526518.

Zelealem Leyew. 2003. Some points on Cushitic and Omotic numerals. Presented at the 4th International Conference of Cushitic and Omotic Languages, April 10–12, 2003, Leiden University, The Netherlands.

Zorc, R. David & Madina M. Osman. 1993. *Somali-English dictionary with English index*. Kensington: Dunwoody.

DOI 10.15460/auue.2025.98.1.368

Appendix: Assessment of the proposed cognate sets

The following overview discusses the thirty-seven lexical items that Heine et al. (1979) set up for Proto Baz as the source for Cushitic loans in South Nilotic. After the proposed word/stem as heading, we provide their proposal, consisting of a correspondence set of Proto Baz, Proto South Nilotic/Proto Kalenjin/Common Datooga and Proto East Omo-Tana. A second etymon is provided from later reconstructions (Rottland (1982) for South Nilotic, Heine (1978, 1981) or Lamberti (1986) for East Omo-Tana) if it differs from the one in Heine et al. (1979). Next, cognates (or related items) in relevant Cushitic and other languages are provided. Each item is concluded with a brief discussion and our verdict. We also include the proposals that Heine et al. (1979) present in endnotes as suggestions from Christopher Ehret in personal communication; here indicated as endnote 29 (for loans from Oromoid), 32 (for loans from Yaaku-Dullay) and 41 (for loans from East Cushitic). Some other Cushitic loans in Nilotic that Ehret proposed in his PhD (Ehret 1970, 1971) are added and marked as B1 for Table 1 (East Cushitic) in Ehret (1970 or 1971).

For ease of reference, we have ordered all items alphabetically on the abstract title of the items in capital letters. Please also note that Heine et al. (1979) use “PS” for Proto Sam, which we instead refer to as Proto East Omo-Tana; this means it is abbreviated with “PS” in the proposed Proto Baz lexical item line but referred to in text as “Proto East Omo-Tana”.

Abbreviations

CD Common Datooga
PB Proto Baz
PEC Proto East Cushitic
PK Proto Kalenjin
PS Proto Sam
PSN Proto South Nilotic

DOI 10.15460/auue.2025.98.1.368

AFARTAM ‘forty’ and other even decimals

Heine et al. (1979) include *afartam ‘forty’ in their inventory of Proto Baz words, but do not discuss it in the body of the article, nor do they include *f* in their inventory of Proto Baz phonemes. On the basis of the fact that Nandi has *artam* ‘40’ (Creider & Creider 2001) and Datooga-Asimjeega *ardam* ‘40’ (Griscom ed. 2018: 9), we can assume a transfer into Proto South Nilotic *artam from East Omo-Tana *afartam (extrapolated from Somali *afartan*) with loss of intervocalic *f*, from a higher Cushitic level.

Ehret (1970: 146; 1971: 112) proposes Proto South Nilotic *tɪktɛm ‘20’ to be from Cushitic, e.g. Oromo *digdama* (confirmed in Gragg 1982) with *tɛm* from Cushitic *tomon* ‘ten’. Rottland (1982: 563) does not reconstruct it at Proto South Nilotic level but has Proto Kalenjin *tiptem and Common Datooga *digdam*. Oromo is the only Cushitic language that has such a word for ‘20’.[4] Other Oromoid languages have constructions which are likely to be lexical innovations. Highland East Cushitic languages have a word for 20 based on *lam-* ‘two’ such as Burji *lamat(t)an*, as do the Agaw languages *läŋäːtäŋən* with gloss /two:ten/ (Appleyard 2006: 140).

Ehret (1980: 147–8; 1981: 112) adds Datooga *laːdʌm* ‘60’, *ladám* in Griscom (ed. 2018: 28) and *seːdʌm* ‘80’, *sedám* in Griscom (ed. 2018: 43), as tokens of transfer from Cushitic. Somali has *lixdan* ‘60’. We support these proposals of transfer of ‘20’, ‘40’, ‘60’, and ‘80’ from a Cushitic source to proto South Nilotic.

AM

- PB *am ‘to eat’ : PK *ʌm : Proto-East-Cushitic *kom (Sasse 1979: 25)

Rottland (1982: 296) reconstructs this item to Proto South Nilotic as *am-it ‘food, sustenance’ on the basis of Proto Kalenjin and Datooga *ham* (n) ‘food’; we propose to adjust the reconstruction to *ham-it to accommodate for Datooga *ham* and in line with the Omo-Tana attestations. Lamberti (1986: 298) reconstructs *aham for Proto East Omo-Tana. The Proto East Cushitic form *k’om ‘chew, bite, eat’ presented in Heine et al. (1979) is not cognate with Proto East Omo-Tana *aham, because its regular cognate in Somali is *qoom* ‘to wound’ (Sasse 1979: 25). Within Cushitic the relevant root is present in West

4 Also Gedeo *diddama,* but Gedeo has borrowed higher numerals from Oromo.

DOI 10.15460/auue.2025.98.1.368

Omo-Tana: Arbore *ʔɔhɔm* ‘eat’ (Hayward 1984: 417), Elmolo: *aam* ‘eat!’ (Tosco 2015: 117) and is reconstructable for Omo-Tana. This item then could represent transfer from East Omo-Tana or Omo-Tana to Proto South Nilotic.

Alternatively, the similarity of these words can easily be due to their sound-symbolic value. For example, Somali has the ideophone *ham* to indicate an action of eating greedily, wolf down (Dhoorre & Tosco 1998: 135, 144).

APIYE ‘animal’ (B1)

Ehret (1971: 111) has *apiiye* ‘large carnivore’ as Proto South Nilotic reconstruction and mentions *apiːrie* as Pre-Proto South Nilotic form. No evidence is presented but note that Pre-Proto South Nilotic *R indeed continues as *y* in most Kalenjin varieties in Rottland’s reconstruction. Ehret links the root to Oromo *arba* ‘elephant’ and Yaaku *arape* ‘large feline’. The root is cognate to Turkana *epiri/ngipiryo* ‘hippopotamus’ (Ohta 1989: 33) and has spread to Ik *jnépírìà* ‘hippo’ (Schrock 2017: 195).

However, Rottland (1982: 465) has only Datooga CD *(h)abiy* ‘hyena’ and no reconstruction for Proto South Nilotic. The semantic link between the Cushitic root for ‘elephant’ and the Datooga word for ‘hyena’ is problematic. We cannot imagine a scenario in which a word meaning ‘elephant’ is interpreted as meaning ‘hyena’. In our view this is not solved by proposing a general meaning such as ‘large carnivore’ as we are unaware of this being an attested lexical semantic unit in the languages of the wider area, nor is the elephant a carnivore.

The root may be related to Proto East Cushitic *ʕarab- ‘elephant’ based on: “Somali *arb-e* (Andrzejewski), Rendille *arab*, Elmolo *arap*, Arbore *arb-a*, Dasenech *’arab, pl. *’arb-u; Galla *arb-a*, Konso *arb-a*, Gidole *arp(-a)*; Burji *arb-a*; Harso and Gollango *arap-ko*; Yaaku *arap-e* ‘large feline’ (Ehret), *arap-a* ‘carnivorous animal’ (Heine).” (Sasse 1979: 14).

The presence of the root meaning ‘hippo’ in Teso-Turkana (East Nilotic) indicates that the Nilotic distribution of a possible reconstruction is wider or higher than Datooga though. It is not reconstructed for Proto Nilotic in Dimmendaal (1988). The semantic link between elephant and hippo is easier to understand than from those meanings to ‘hyena’. We are not yet convinced of transfer from a

DOI 10.15460/auue.2025.98.1.368

Cushitic source to South Nilotic because of the lack of evidence in Kalenjin and the semantic distance between 'elephant' and 'hyena' but we consider the link between an East Cushitic word for 'elephant' and the words for hippo in East Nilotic Turkana (and Ik) promising.

AR

- PB *aR 'male bovine' : PK *e:ɣ : Dhaasanac *ar*, Arbore *ar*, Bayso *are*

Rottland (1986: 297, 556) reconstructs *e:R to Proto South Nilotic, as well as to Proto Kalenjin. Lamberti (1986: 215) mentions an item similar in form and meaning for Central-North Somali, Rendille, Jiiddu and Boni, with Central-North Somali *awr* 'male camel' being the most archaic form in that it preserves the diphthong; Lamberti does not, however, reconstruct this item to Proto East Omo-Tana for unknown reasons. Rendille *oor* 'bull of cattle or camels' is supposedly related but differs in having a round vowel.

The closeness of Arbore *ʕáar* 'bull' and Dhaasanac *ar* 'bull' in form and meaning to the Proto Kalenjin form *e:R suggests the possibility of a transfer from West Omo-Tana (Dhaasanac) to Nilotic. This is likely to be a transfer from Proto West Omo-Tana into Proto South Nilotic assuming centralisation raising due to the R sound and no need to explain the loss of roundness if it had been borrowed from East Omo-Tana.

Datooga æ:nog 'milk' from East Cushitic (B1)

This is proposed in Ehret (1981: 110) as an East Cushitic loan in South Nilotic providing Oromo: *ana:n*, Somali *ʕaano*, Saho *hān-* as support. Rottland (1982: 547) has POD *a:n; and Common Datooga: *a:n* and these are closer in form to the East Cushitic forms: (*-g* is a number suffix). Black (1974: 219) reconstructs Proto Lowland East Cushitic $ʕ_1$aan- on the basis of Afar *ħan* and Proto South Lowland East Cushitic *ʕaan-. Heine (1978: 77) has Proto East Omo-Tana *ʕaanu, Sasse (1979, 1982) does not reconstruct it. Transfer from a Cushitic source (Lowland East Cushitic, Omo-Tana or East Omo-Tana) into Datooga is likely but it is strange that the transfer is to Datooga and not to Proto Kalenjin or Proto South Nilotic.

DOI 10.15460/auue.2025.98.1.368

BAZ ‘lake’

- PB *baz ‘lake, sea’ : Datooga bas- (< *pas-) : PS *be

Common Datooga has a slightly different vowel in Rottland (1982: 554): *bas* ‘lake, sea’. Beyond South Nilotic, this item is reconstructed to Proto Nilotic *bas ‘lake, wide river’ (Dimmendaal 1988: 46), cf. Asimjeeg Datooga *bàsòːd* (Griscom 2019: 288), Turkana *em-baso* (Beech 1911: 137), Maa *básu* (*m-*) ‘lake’, *pásô* (*ém-*) ‘lake, large pool’, and Proto-West Nilotic *bar ‘river, lake’. There is no evidence for this root in Kalenjin (Nandi, Pokot, Marakwet, Sabaot), not even as a name for Lake Turkana (Pokot).

The word is wide-spread in Cushitic with reflexes found in various branches: Proto East Omo-Tana: *bez (f) from Rendille *bey*, pl. *bey’ay* (f/m), Somali *bad* (f). But also at the higher Proto East Cushitic level, it has been reconstructed by Sasse (1979: 20) as *baz- ‘lake, sea’, based on Saho-Afar *bad* ‘sea’; Somali *bad* ‘sea’, Rendille *bey* ‘lake’, Dhaasanac *baz* ‘Lake Rudolf’, Elmolo *paw* ‘Lake Rudolf’; Gollango and Dobase *pas-o* [*baz-o*] ‘lake’. The reverse direction, Proto Nilotic into Proto East Cushitic would leave *z in Cushitic unexplained (Proto Nilotic has *s); hence Proto East Cushitic into Proto Nilotic is more likely.

Dhaasanac is the closest formal and geographical source among the attested or reconstructed Cushitic candidates. The Proto Nilotic item however necessitates a source at a deeper time-depth, unless the word was a later transfer that spread within Nilotic in a fashion of etymological nativisation, with the result that the reconstruction is an ahistorical mirage. At this point, however, there is nothing to suggest this. We propose that this is an early loan from Cushitic to Nilotic.

BOD ‘get lost’

- PB *bod ‘to get lost’ : Pokot *pot* : Proto-East-Cushitic *bad

There is no Proto East Omo-Tana reconstruction for this item, but the root is widely attested in East Cushitic: Proto East Cushitic: *bad- ‘be lost, extinguished’: Saho-Afar *ba(a)d-* ‘perish, be extinguished’; Dhaasanac *bad-*, Elmolo *pɛd-*; Oromo *bad-* ‘be lost’; Konso *pat-* ‘be ruined’; Dirayta *pat-* ‘disappear, be lost, die’; Burji *bas-s-* (caus.) ‘extinguish’ (< *bad-s-); Gawwada and Harso *pot-* ‘disappear’, Gollango *pot-* ‘get lost’. Probably Somali *bad-* ‘involve in trouble’ and *bad-baad-* ‘escape’ also belong here.

DOI 10.15460/auue.2025.98.1.368

The Dullay form is closest to the Pokot form and one can suggest a Dullay to Pokot transfer; not excluding other Cushitic sources as options because both Dullay and Kalenjin do not distinguish between voiced and voiceless stops and the realisation of *a* and *o* is close; hence the similarity between Dullay and Pokot may be not markedly closer than with other possible sources.

Datooga has *bàd-éeda* 'trouble' which suggests PSN *bad 'get lost' from PEC as pointed out to us by an anonymous reviewer. The direction of transfer can easily be the other way given presence of this root in Majang *bàdɛ́ɛ́r* 'disappear, get lost' and Ik (Kuliak) *bàd-* 'die'.

BOQOL 'hundred'

- PB *boqol 'hundred' : PK *pOkOl , Datooga *boqal* : PS *boqol

Rottland (1982) reconstructed Proto Kalenjin *pɔkɔl and Common Datooga *boqal*. It was not reconstructed by Rottland for Proto South Nilotic but this may be possible. There is the issue of Proto South Nilotic *q; Dimmendaal (1988: 7, 10) tentatively reconstructs this phoneme for Proto Nilotic, at least before low back vowels. In this view, we would tentatively reconstruct Proto South Nilotic *poqol with a development $q > k$ in the transition to Proto Kalenjin. This opens the scenario for a transfer from Proto East Omo-Tana to Proto South Nilotic. The root is absent in Cushitic outside of East Omo-Tana and the other direction of transfer cannot be excluded. We propose the direction into Proto South Nilotic mainly to let it be in line with the transfer of the other higher numerals from Cushitic into Proto South Nilotic.

BURI red

- PB *buri- 'red' : PK *pirir : Dasenech *bur*
- Rottland (1982: 308) PSN piriir 'red'

This item is only attested in West Omo-Tana within Cushitic so far, with Elmolo *poor* 'red', Arbore *burrí* (*rr* is a trill – while *r* is a flap in Arbore), Dhaasanac *bur* 'red' (Tosco 2001). Proto West Omo-Tana *buri is proposed.

Obstruents in South Nilotic became devoiced as there was no voice opposition in obstruents in South Nilotic while in Proto West Omo-Tana initial bilabial stops were voiced (Hayward 1984: 33). Proto South Nilotic has three consonants in this root. If Proto South Nilotic

DOI 10.15460/auue.2025.98.1.368

borrowed from Proto West Omo-Tana it either reduplicated the final consonant, or it inserted a vowel in the geminate final *r*. If Proto West Omo-Tana borrowed from Proto South Nilotic, it lost the vowel of the second syllable and changed the vowel *i* to *u* under influence of the initial labial consonant. The direction Proto South Nilotic to West Omo-Tana is easier to explain. Moreover, the root seems to be isolated within Cushitic.

ɗERI ‘clay pot’

- PB *d’eri ‘clay pot’ : PK *ter(et) : PS *d’eri (= Heine 1978)

Rottland (1982: 441) reconstructs Proto Kalenjin *ter ‘clay pot, pot’. This item seems to have a restricted distribution both in South Nilotic (only Kalenjin) and East Cushitic (not in Highland East Cushitic, Proto Lowland East Cushitic, Oromo). *ɗ* adapted as *t* is more easily explained than *t* > *ɗ*. This suggests transfer from East Omo-Tana to Proto Kalenjin.

PK *eu(n) ‘night’

Heine et al. (1979: endnote 32) report that Ehret claims this root to be among the Kalenjin words from Yaaku-Dullay. Ehret (1970, 1971) does not mention this root for ‘night’. Rottland (1982: 548) has Common Datooga *ɛw ‘night’ and does not reconstruct it for Kalenjin, nor the higher South Nilotic level. Nandi has no potential cognate for ‘night’ and ‘evening’ (Creider & Creider 2001); and nothing has been found in Pokot, so it cannot be shown to be a transfer at Proto South Nilotic level rather than to Datooga.

The Datooga root is likely to be a reflex of a Cushitic transfer into South Nilotic. Proto Yaaku-Dullay is the most likely source: Gawwada *ʔawn-e*, Tsamakko *ʔawne*, Yaaku *aun*; Ehret (2008) suggests *ʔawn ‘night’ for Proto Yaaku-Dullay). Other cognates in Cushitic such as Oromo *ʔuo* and Rendille *iben* differ more in form. The receiving language lost the final *-n*. If the receiving language was (Proto) Datooga, there is the option that Datooga speakers interpreted the final *-n* as their singulative suffix *-an*. There is enough evidence to suggest a Yaaku-Dullay borrowing into Proto Datooga.

GID ‘jigger’ (B1)

Ehret (1970: 148, 1971: 111) links Datooga *gidgidood* ‘jigger’ to Somali *kudkude* and Saho: *kudkud* ‘tick’ (not in Banti & Vergari 2005

DOI 10.15460/auue.2025.98.1.368

or Vergari & Vergari 2007). Similar items can be found in Proto Highland East Cushtic: *kora 'flea' (Hudson 1989) and Oromo *tuffee-kurup'p'ee* 'flea' (cf. *kurup'p'isa* 'grasshopper') (Gragg 1982: 379, 256, 457: *tuffee-* 'any thing or anyone that makes a sudden appearance, penis, anus'). Rottland (1982: 537) has Proto Kalenjin *kɪmɪt* 'flea'. The evidence is not strong enough to propose a transfer. See INSIR 'louse' below.

GORO 'feathers, fur'

- PB *goro 'fur' : PK *karaːr : Bayso *ogoro*

This item is not reconstructed to Proto East Omo-Tana by Heine (1978, 1981) nor Lamberti (1986). It may be compared to Tsamakko *koolo* 'wing', Gawwada: *hool-o* (m) 'wing, feather', and Konso: *xool-a* (m) 'feather (including down), wing'.

Rottland (1982: 334) reconstructs indeed Proto Kalenjin *karaːr 'feather, hairs (on the fur)'.

In conclusion, there is no clearly related item in Cushitic. The best source candidate presumably is Dullay *koolo 'feather, wing' (assuming Dullay to Konso transfer in this instance). If it is related, the final *r* in Proto Kalenjin is unexplained. Pending further evidence, this item does not provide strong support for transfer from Cushitic into Proto Kalenjin.

HAB

- PB *hab 'curse' : Datooga *(h)abɛ : PS *hàbáàr

The Datooga form is not found in Rottland (1982) nor in Griscom (2019). There is no other evidence in South Nilotic. In Cushitic, beyond East Omo-Tana, this item is also found in Oromo *abbaara* 'curse, scold'. We cannot draw any conclusions given the lack of evidence for the item in Datooga, or in South Nilotic.

Swahili *-apa* 'swear' is unrelated despite its similarity in form and meaning. It is in some dialects still *-lapa* (Sacleux 1939) and it is a regular reflex of Proto Bantu *dap 'swear' (see BLR3 #872).

INSIR 'louse'

- PB *insir 'louse' : PK *(k)ɪnsɪr : PS : *ingir-

Rottland (1982: 545, 348) presents a finer-grained picture of Proto Kalenjin: *(k)ɪ(n)sɪr, with *-n-* being found only in Pokot. The loss of

DOI 10.15460/auue.2025.98.1.368

n in nearly all languages is irregular within Kalenjin. We suggest that Pokot borrowed *ɪnsɪrya:n* from Oromo *injiraan*; Pokot adjusted the *j* sound to their phonological system. We adjust the Proto Kalenjin reconstruction to *(k)ɪsɪr without *n*. Lamberti (1986: 364) reconstructs Proto East Omo-Tana *ingir-, presumably with an eye on the velar in the reflexes in Elmolo and Arbore but this would fit a higher level reconstruction, Proto Omo-Tana, but not at the lower East-Omo-Tana level since none of the reflexes in East Omo-Tana have *g*. A reconstruction *injir for East Omo-Tana better fits the East Omo-Tana evidence and is a more likely source for Proto Kalenjin: *(k)ɪsɪr.

The other Nilotic branches have different roots for 'louse'. Dimmendaal (1988) reconstructs Proto Nilotic *ɲok. On this basis we can assume that Proto Kalenjin borrowed the item and was not the source. Rottland (1982: 347–8) compares the addition of an initial *k*- in Keyo, Tuken and Sapiny to the addition of *k* in *kitook* 'place to sleep' but cannot explain it. Possibly the shape of the word of 'flea' *kɪmɪt* played a role. If we suppose a transfer from Cushitic to Kalenjin *ɪsɪr*, the closest source in shape would be East Omo-Tana *injir*. We still need to explain the interpretation of *nj* as *s*. The devoicing of *j* can be explained by the absence of phonological opposition in voice in obstruents in Kalenjin. The loss of *n* suggests a source that realised the initial *in* as a nasal vowel.

Dhaasanac has a deviant shape for this root within Cushitic, *iðír-ri* 'louse, lice' (Tosco 2001: 482)[5] (different from related Arbore *íngir-a* and Elmolo *íŋkir*) and the Dhaasanac form may have been a transfer from a Kalenjin source with *isir*, assuming adaptation of intervocalic *s* to *ð*.

It appears that Kalenjin languages added an initial *k*- in this root; this is difficult to explain. Rottland (1982) does notice a parallel case with an unstable initial *k*- but also remarks that he cannot explain it. Transfer from East Omo-Tana into Proto Kalenjin is conceivable.

ILAAL 'look'

- PB *(i)laal 'to look' : PK * (i)la:l : PS *ilaal

Rottland (1982: 552, 554) does not reconstruct this item for Proto Kalenjin, and this means that at present there is no evidence of transfer from Cushitic into Proto Kalenjin for this item. Ehret (1971:

5 *izid-dî* in Black (1974: 230, 299) is the singulative.

DOI 10.15460/auue.2025.98.1.368

110) reconstructs Proto Kalenjin *(i)laal ‘to watch, to look’ and links it to Oromo *ilaal-* ‘to look at’, Somali *ilaali* and Saho *ilal-* ‘to look’ and suggests a Cushitic internal derivation from *ila ‘eye’. On this basis he proposes a transfer from East Cushitic into South Nilotic. He does not provide evidence for his Kalenjin reconstruction.

IRI ‘smoke’

- PB: no evidence.

Heine et al. (1979) do not provide a Proto Baz or Proto East Omo-Tana form for this item, which only figures in their Appendix (Heine et al. 1979: 86). As they note in endnote 41: “Some roots like ‘calf’ MAR and ‘smoke’ IRI occur also in East Sudanic groups and may therefore not constitute borrowings from Proto Baz. They do, however, fit into the regular sound pattern of borrowing established here and therefore have been tentatively included.” Rottland (1982: 550) reconstructs *i:R ‘smoke’ for Proto Kalenjin. Potentially related forms have not been found in Cushitic; thus we see no evidence for transfer into Proto Kalenjin for this item.

KAS ‘see’

- PB *kas ‘to see’ : PK *kʌs : PS *kas ‘to understand, know’
- PS: Somali: *kas-*; Boni: *kâs/kása*.

Proto Kalenjin *kas* ‘to hear, understand’ (in Sapiny and Bong’om: ‘to hear, see, understand’) (Rottland 1982: 340). Note that ‘to see’ is *da-* in Common Datooga) (Rottland 1982: 554). In Somali, *kas-* means ‘to understand, comprehend; know, have knowledge of’ (Said 2013: 212). While it is attested in Boni, Lamberti (1986) does not reconstruct this item to Proto East Omo-Tana. Semantically, the shift from ‘see’ to ‘know, understand’ is more likely than the other way round, as argued by Sweetser (2019: 38) for the body to mind metaphor sight to knowledge, mental vision. Our conclusion is that if the Proto Kalenjin and East Omo-Tana items are related, it was transferred into East Omo-Tana from Proto Kalenjin.

PK *kolkoc, *kockoc ‘tortoise’ : Oromo qoca

Heine et al. (1979: endnote 29) mention this as a proposal by Ehret for Oromoid influence on Proto Kalenjin on the basis of Oromo *qoca*, which is *qoc’aa* ‘turtle’ in Gragg (1982). It is not attested in the other Oromoid languages: The root is not attested in Dirayta, while Mosiye

DOI 10.15460/auue.2025.98.1.368

has a different root, but the root is reconstructed for Proto Highland East Cushitic as *k'oc'a 'turtle, tortoise' (Hudson 1989: 157). There are no attestations in West- nor East Omo-Tana.[6] The Proto Kalenjin reconstruction cannot be confirmed so far, as it is not included in Rottland (1982) for Proto South Nilotic, Proto Kalenjin, or Common Datooga. Compare, however, Nandi *ceːp-koykóc-êːt* 'tortoise' (Creider & Creider 2001: 64), Päkot *kokech* (Beech 1911: 145), as well as Samburu (East Nilotic) *loi[guiguy]ari* 'tortoise' (Webonary 2019) that are close to Heine et al.'s (1979) Proto Kalenjin reconstruction. Once the Proto Kalenjin reconstruction can be confirmed, this item can be proposed to be a transfer from Proto Oromoid to Proto Kalenjin and the latter must have applied reduplication. We can postulate a Proto Oromoid form *qoc'a given the cognate in Proto Highland East Cushitic and this transferred into Kalenjin, possibly Proto Kalenjin, pending the Proto Kalenjin reconstruction.

KONOM 'fifty' and other uneven decimals

- PB *konom 'fifty' : PK *kɔnɔm : PS *kontom, Somali *konton*

Rottland (1982: 538) reconstructs Proto Kalenjin *kanam with *a*~ɔ in Pokot (Rottland 1982: 334). There is no East Omo-Tana reconstruction for '50' in Heine (1978); the reconstruction in Heine et al. (1979) is probably extrapolated on the basis of Somali. This is a likely transfer from Proto East Omo-Tana or higher into Proto Kalenjin or Proto South Nilotic. Datooga utilises a construction for the decimals which is why there is no higher reconstruction at Proto South Nilotic level for the uneven decimals. All uneven decimals in Datooga are constructions rather than single roots: 90 = 80 + 10, 50 = 40 + 10, 30 = 20 + 10. This construction is an innovation at least in that it uses borrowed roots for '80', '40' etc., but it is very well possible that the twenty-base system is old in (South) Nilotic and that the construction as such in Datooga is a remnant of that. In that scenario the even decimals were borrowed in South Nilotic but the uneven decimals only in Kalenjin.[7]

6 Somali *qubo* 'turtle, tortoise' (Zorc & Osman 1993: 334) and Konso *kup-aata* 'tortoise, turtle; short person' (Black & Otto 1973: 78) are different roots.

7 We thank Thilo Schadeberg for pointing this out to us.

DOI 10.15460/auue.2025.98.1.368

KOR ‘bell’

- Heine et al. (1979): PB *kor ‘bell’ : PK *kʊr : PS *kor-

The reconstruction for Proto Kalenjin in Heine et al. (1979) is apparently no longer considered viable in Rottland (1982). Rottland (1982: 450) has a different root, *twaːl* ‘(cow) bell’, for Proto Kalenjin (cf. TUAALIO below in the Appendix). Oromo and Konso have *korkor* as an ideophone for the hollow sound of something like a camel bell. This word may have circulated widely with the cultural artefact of bells/rattles, and/or it may be sound-symbolic. Cf. also Ik: *coor* ‘leg bangle’, Elgon: *coːr* ‘dance (v)’, Nandi-Markweta: *kurkur* ‘bells on leg’, Mursi-Suri: *joré* ‘decorative bell’, Dholuo: *gara* ‘ankle bells’, Hamar: *c’onc’oro* ‘rattle’. In the absence of any evidence for this root in Kalenjin and assuming that *twɑːl is a different root, there is no indication for transfer from Cushitic into Nilotic.

LVħ ‘six’

- PB *lVħ ‘six’ : PK *ila : PS *liħ

Rottland (1982) has reconstructions with slightly different vowels: Proto Kalenjin *lɔ, Proto Kalenjin-Omotik *lɑ, and Common Datooga *la*. The word is not reconstructed for Proto South Nilotic but the words in the two branches are related, showing irregularities. Sasse (1979: 22, 63) reconstructs *liħ for Proto East Cushitic, as in East Omo-Tana. Proto West Omo-Tana has probably the same proto form, probably with a change of final *ħ* to *h*, even if the individual languages developed different shapes: Dhaasanac *li*, Arbore *ʤih* (Hayward 1984: 439), Elmolo: *yii?* (Heine 1980, regular reflex according to Sasse 1979: 22). The same is valid for Proto Oromoid: Oromo *jah-a*, *jaa* ‘six’ (regular reflex of *liħ), Konso *leh*, Dirayta *leh(-e)*. The Dullay languages have a different root for ‘six’, Gawwada *tappi*; Tsamakko *tabben*, and Yaaku borrowed the item from East Nilotic. The most likely Cushitic source is Proto East Omo-Tana or higher levels such as Proto Omo-Tana or Proto Lowland East Cushitic. Lowland East Cushitic *e* > *i* / _ pharyngeal that is not followed by C (Black 1974: 111). So, unless the pharyngeal was followed by a C in ‘six’, then this item is irregular in Black’s system. Sasse (1979: 22) solves it by positing a variant form at a higher level: Oromo, Baiso, Konso, Gidole developed from “a variant *leħ”.

DOI 10.15460/auue.2025.98.1.368

Proto Highland East Cushitic has *leho* (Hudson 1989: 135). Proto-Agaw has a different root *walta (Appleyard 2006: 124) and Beja has *asagwir* (Wedekind et al. 2021). But Tanzanian Cushitic has a cognate root closest to Proto Highland East Cushitic, *laħoo?u (Kießling & Mous 2003: 190), suggesting Pre-East Cushitic *laħo [leħo] > Proto Highland East Cushitic *leho; parallel > Lowland East Cushitic [leħo] (as it is in Oromoid, see Sasse (1979: 22) > Proto East Omo-Tana *liħ (no explanation for this raising); and parallel (or earlier) > Proto West Rift *laħo-?u. The indication *laħo [leħo] refers to the situation that a phonological /a/ is pronounced as [e] before a pharyngeal.

The receiving South Nilotic languages lost the final pharyngeal fricative which is absent in their inventory and interpreted the vowel which was allophonically centralised [ɒ] before the pharyngeal in the Cushitic source as a lower centralised vowel. The reconstructed rounding of the vowel in Kalenjin (Rottland 1982: 366) is difficult to explain but all Kalenjin lects have a round vowel. Datooga has a low vowel in all lects and the *l* is often long. The extra initial vowel *i* as it appears in Heine et al. (1979) is not retained in Rottland (1982), but it is present in the East Nilotic languages that borrowed this root: Proto-Teso-Masaian has *ille (Vossen 1982: 99 going back to Ehret 1974: 40). Blažek (2018: 50) also includes a syllable before *liħ*: Proto East Cushitic *(ʔ/ha)liħ. It is difficult to determine how many transfers happened and between which proto languages.

MALAB ‘honey’

- PB *malab ‘honey’ : Datooga *mal* : PS *málàb

This item is widely attested in East Cushitic and is reconstructed to Proto East Cushitic: *malab- ‘honey’: Saho-Afar *mala(a)b-*; Somali *malab*, Rendille *malab*, Boni *malub*; Sidamo and Kambata *malab-o*, Hadiyya *marab-o* (Sasse 1979: 14). Arbore *múul* fits Datooga in terms of syllable structure, but not the vowel. Transfer into Datooga is possible, maybe from a language that lost the rhyme of the second syllable, like Arbore. In that case it would have taken place before the change of *a* > *u*. The best-supported scenario is transfer from West Omo-Tana into Datooga.

MAQAL ‘young goat/sheep’

- PB *maqal ‘young goat, sheep’ : PK *mʌkʌl ‘male sheep’ : PS *maqal ‘young goat, sheep’

DOI 10.15460/auue.2025.98.1.368

Rottland (1982: 369) does not contain this item, but it is found in Pokot *màkáal* 'ram' (Crazzolara 1978: 239). Ehret (1970: 138) has Proto Kalenjin *makal 'male sheep' but does not propose Cushitic influence. Within Cushitic, the root seems to be restricted to East Omo-Tana. The most likely scenario is transfer from East Omo-Tana into Pokot.

MAR 'calf'

- PB *maR 'calf' : PK *mɔːɣ : Rendille *mar*

Rottland (1982) reconstructs *mɔːR* at the higher Proto South Nilotic level (Rottland 1982: 304, 377, 543). Pillinger & Galboran (1999) have for Rendille *maár* 'female calf', *máar* 'male calf', and this long vowel is a better fit with Proto Kalenjin than Proto Baz *mar. The item is not present in Proto East Omo-Tana (Lamberti 1986), but in the West Omo-Tana languages Elmolo (*maar* 'calf'), Arbore (*máar* 'calf') and Dhaasanac (*moɗo* 'calf').

The item is found in East Sudanic beyond South Nilotic, e.g. Mursi-Suri *mòr* 'female calf'. It cannot be excluded that the direction of borrowing was from South Nilotic to West Omo-Tana. In fact, given that the item is limited to West Omo-Tana in Cushitic, the possibly irregular *ɗ* in Dhaasanac, and the occurrence in Mursi-Suri, it is more likely that the item passed into Cushitic from South Nilotic/Surmic. It is also widely attested in Bantu languages of the region. Ehret (1970: 60) "One piece of evidence does suggest an area for possible further investigation: the word *moori*, "calf" or "heifer", related to Proto Southern Nilotic *mɔɔi < Pre-Southern Nilotic *mɔɔri, "calf", is found not only in all East Victoria Bantu dialects and in Thagicu languages, but also in Shambala, Zigula, and Swahili. But if Southern Nilotes did indeed directly influence the ancestors of these three Bantu peoples or others among the far northeastern Bantu, that influence is not easily detectable in their languages." If Cushitic did not borrow it from East Sudanic, South Nilotic may have taken it from Rendille or West Omo-Tana.

MEE 'die'

Heine et al. (1979: 88 endnote 32) report that Ehret proposed to them that Kalenjin *me: is from Yaaku-Dullay origin. Rottland (1982: 556) has Proto South Nilotic/Proto Kalenjin *mɛːR* 'to die'. The root can possibly be related to Yaaku *mɛʔɛ* 'to get lost (of animals)' but no

DOI 10.15460/auue.2025.98.1.368

other potential cognate was found in Dullay, in Gawwada nor Tsamai. Given the weakness of the semantic fit and the lack of additional support, we consider this not to be a Cushitic transfer into South Nilotic.

METEH 'head'

- PB *meteħ 'head' : PK *meh : PS *mataħ

This item has been reconstructed to Proto East Cushitic: *math- 'head' and is widely attested: Somali *madaħ* Baiso *mete*, Rendille *mataħ*, Boni *mada?*, Arbore *mete*, Dhaasanac *mé* (irregular loss of *t*; cf. pl. *mett-u*) (Tosco 2001: 517), Elmolo *mete?*; Oromo *mataa*, Konso *matta*, Dirasha *mašš(a)*; Yaaku *miteh* (Sasse 1979: 10). The closest match to Proto Kalenjin is Dhaasanac. If the Proto Kalenjin item is a loan, its likely source is Dhaasanac. But Rottland (1982: 302) has Proto South Nilotic *met 'head'. Hence the source can be any level of Cushitic and the receiving language lost the final *ħ* or *h*. West Omo-Tana fits best for the vowel *e*.

MIE 'good' (B1)

Rottland (1982: 540) has Common Datooga *miiy* 'good' but does not reconstruct it to Proto South Nilotic level. Ehret (1970: 148; 1981: 110) links Proto South Nilotic *miie* 'good' to Saho *maʕa* 'good', Oromo *mia* and Somali *maʕaan* 'sweet'. The Oromo word has different semantics referring to things rather than quality, and is not linked; Gragg (1982) has *miʔa* (subj. *miini*) (n) 'goods, possessions, things'. The Saho word is confirmed by Banti & Vergari (2005: 16) as *maʕaani* (f) 'goodness, kindness, righteousness'.

The vowels in the Cushitic root can be considered to be centralised due to the presence of the pharyngeal; raising to *i* is more difficult to explain. All in all, the evidence is limited, but an etymological link cannot be ruled out. It is not clear from which Cushitic group to which Nilotic group the transfer would have occurred.

MOOSOONG 'sorghum'

Ehret (1970: 146; 1971: 111) links the Kalenjin root *moːsoːŋ to Sidamo *meshinga* (Rottland (1982: 542) has the same Proto Kalenjin reconstruction *moːsoːŋ). However, Proto Highland East Cushitic is *bashink'a (Hudson 1989) and Yri & Pepper (2019) have *bashaank'a* 'millet' for Sidaama (= Sidamo). Ehret mentions Oromo *misinga*, which matches Orma *misingaa* 'millet' (Stroomer 2001: 121), cf. Wel-

DOI 10.15460/auue.2025.98.1.368

legga Oromo *bisinga* ‘sorghum’ (Gragg 1982). There are known cases of alternations between *b* and *m* in East Cushitic and Ethio-Semitic, but the dissimilar vowels *o* and *i* in Proto Kalenjin and Orma do not suggest a direct link between Proto Kalenjin and Orma here. However, the Proto Kalenjin item may well be an instantiation of a *Wanderwort* in the area (cf. Amharic *mashilla* ‘sorghum’), without necessarily being a transfer from Cushitic.

MUR ‘cut, circumcise’ (B1)

Ehret (1971: 146) has Proto Kalenjin *mʊraataan ‘circumcise’ (similarly, Proto Kalenjin *mʊrɑːtɑn ‘circumcise’ in Rottland (1982)), *muren* ‘warrior’; and he links this to Oromo and Sidamo, East Cushitic, *mur-* ‘to cut’. Rottland (1982) has Proto Kalenjin *muren for ‘man, circumcised man’. The Kalenjin root for circumcised man may very well be related to Proto East Cushitic *mur- ‘cut, judge’ as reconstructed by Sasse (1979) on the basis of Dhaasanac *mur-*, Elmolo and Arbore *mur-i-* ‘short’; Galla *mur-*, Gidole *mur-* ‘cut person’s genitals to take trophy’, Konso *mur-*; Highland East Cushitic (Sidamo, Haddiya) *mur-*; Burji *mur-*. Perhaps Dullay *murr-* ‘pay’ is also cognate. There are no attestations in East Omo-Tana. The root is absent in South Cushitic.

Given the data at the Cushitic side, the Cushitic source can be at any historical level of East Cushitic. Ehret (1970) suggests an earlier form in the East Victoria branch of South Nilotic: *mura ‘cut’ (no evidence provided). This would suggest an early transfer and a higher level of Nilotic than Proto Kalenjin. This is confirmed by two reconstructions at Proto Nilotic level by Dimmendaal (1988), i.e. his number 28 *mʊr ‘circumcise, cut’ and number 108 *mʊr ‘marry’ based on *a-mʊran* ‘relation through marriage’ in Teso and *muren* ‘husband’ in Kipsikiis, and additional evidence from Turkana and Pokot with the possibility of a meaning like ‘pass from one stage/age-set to another’. In conclusion, there is early Cushitic influence on Proto-Nilotic and a complex semantic development due to two similar roots with relatable meanings.

There is a similar but probably unrelated root in Proto Kalenjin, *mʊːt ‘to cut, break’ (Rottland 1982: 553, 382). However, all his Kalenjin attestations have a final *t* rather than *r*. He does not reconstruct it up to Proto South Nilotic level due to its absence in Datooga.

DOI 10.15460/auue.2025.98.1.368

In order to reach consistency, we should assume that the Proto Kalenjin item *mʊːt ‘cut, break’ is unrelated.

NUG ‘suck’

- PB *nug ‘to suck’ : Datooga nuq : PS *nuug

Besides in East Omo-Tana, this item is also found in Konso *luuk-* ‘suck’ and other Konsoid varieties (Black 1976: 303). Given the wider distribution in Cushitic, we assume transfer from East Omo-Tana into Datooga, or Proto South Nilotic and later loss in Kalenjin and Omotik. In West-Rift South Cushitic, Iraqw and Gorwaa have a root *nuu-nuu?* ‘suck breast’ which Kießling & Mous (2003) relate to Datooga *nuq*. Two scenarios can be proposed for this link. One is that Pre-Proto-West-Rift South Cushitic had a form *nuq* and this was transferred into Datooga and Datooga does not fit in this cognate set with East-Omo-Tana; or the predecessor of Iraqw and Gorwaa borrowed the word from Datooga in a slightly more specific meaning.

PATAI ‘back’

Ehret (1971: 146) relates his Proto South Nilotic *patai to Somali *baḍi* ‘buttock’ and Sidamo *bade* ‘back’ (*bad’e* in Hudson 1989: 23; not attested in the rest of Highland East Cushitic). Rottland (1982: 551) has Proto South Nilotic/Proto Kalenjin *pataR* ‘back’. Rottland’s *R* is a voiced velar fricative in Pokot, *r* in North Markweta, elsewhere in Kalenjin it is *y*, and in Datooga it is *w/y*. There is no source for this “consonant” on the Cushitic side. The root seems to be restricted to South Nilotic. Somali has indeed *baḍi* ‘buttock’ and Rendille *beḍéy* ‘buttock’. The root is not reconstructed in Heine (1978), but Sosal (in prep.) has Proto East Omo-Tana *baḍey. Potentially, Proto South Nilotic *pɑtɑR transfers to the common ancestor of Oromoid and Omo-Tana (Black’s South East Cushitic), assuming the isolated occurrence in Sidamo is a loan.

QUAR ‘male goat’

- PB: *quar- ‘male goat’ : Datooga *qwaray* : Elmolo *kórat*

Rottland (1982: 357, 362) reconstructs Proto Kalenjin *kwɛR* ‘goat’ with a question mark. In spite of Datooga *qwaray* ‘goat’ (Rottland), Asimjeeg *qwarayd* ‘male goat’ (Griscom 2019: 328), the item is not reconstructed to Proto South Nilotic due to formal incongruencies. This item is found in several branches of East Cushitic, besides

DOI 10.15460/auue.2025.98.1.368

numerous Bantu languages (zones F, G) as well as Kuliak: Burji/Sidamo *kola* 'castrated ram', Elmolo *kor-at* 'goat' (Heine 1980: 207), Dhaasanac *kole* 'male goat', possibly Arbore *k'oll* 'cattle, wealth' (Hayward 1984: 379), Rendille *kelex* 'castrated he-goat', Yaaku *kɔllɛh* 'castrated he-goat', Dirasha *k'ol-d'a* 'goats, general', Ik *kɔla* 'castrated he-goat', So *kol* 'he-goat'.

Rottland states that it is not certain that the forms within Kalenjin are 'direct cognates'; this may point to borrowing after Proto Kalenjin had disintegrated. Breaking of *o* into *wɛ/wa* is not common in Kalenjin; hence it is more likely that this change took place in the donor language. Apart from this vowel, West Omo-Tana, and especially Elmolo, are the closest in form and geography. Proto West Omo-Tana may have been close to *kol- and closest in form (and geography). In line with other transfers from Proto West Omo-Tana to Proto South Nilotic, we can suggest the same for this item leaving the second syllable *-ay* in Datooga to be explained.

QUUT 'scratch'

- PB *quut 'to scratch' : PK *ku:t : PS *qut
- PK *ku:t 'to scrape, to smooth' with meaning 'to weed' for Sapiny (Rottland 1982: 353)
- PS: *qut/quta 'to dig, cultivate'. Rendille *xút; xût/'xúta*, Somali *qod; qood* B *od/óda.*

In terms of form, this could be a transfer from either East Omo-Tana, West Omo-Tana, Proto Omo-Tana or Proto East Cushitic, although the vowel length in Proto Kalenjin needs explanation. Sasse (1979: 10) reconstructs Proto East Cushitic *k'ot- 'dig' adding Arbore *kot-* 'plow', Dhaasanac *g'ot/z-* 'dig, bury'; Oromo *k'ot-*, Konso *qot-*, Dirayta *k'oš-*; and Gawwada/Ale *qot-*. A second obstacle is the semantics, and the presumed direction of the semantic change: 'scratch' to 'dig' is conceivable, while the opposite direction assumed by Heine et al. (1979) is less obvious. In terms of semantics, for most of the (proto-) languages concerned the meaning is both 'to scratch' and 'to "scratch the land"'. The differences in meaning in the proto-languages are only apparent; the polysemy is present in both source and target.

The receiving language is here claimed to be Proto Kalenjin, but (East-)Datooga has *qur-qur* 'dig' (Rottland 1982: 499) which seems cognate. Rottland (1982) does not discuss the development of Datooga q, t; and r is stable. Yet, his PSN *kuut 'blow' is based on

DOI 10.15460/auue.2025.98.1.368

PK *kuut and CD *quur* (Rottland 1982: 300), apparently an irregular series. Moreover, a similar root is reconstructed by Vossen (1982: 417, 236) for East Nilotic at the Proto Teso-Lotuxo-Maa level (= all but the Bari-group) *k_2oɟ- 'scratch', a virtual reconstruction for Eastern Nilotic would be ^{o}kod^{y}. In the absence of regular correspondences, both within South Niltoic and with East Nilotic, Dimmendaal (1988) does not reconstruct this root for proto Nilotic. We can add East-Nilotic Turkana *kut-* 'to dig', *kuta* 'hoe' (Barrett 1988) to the complexity. This root must have been transferred several times. In addition, Taita Bantu has *-kota* 'dig' from Cushitic (Ehret & Nurse 1981: 158).

RAARO 'piece of hide used as container'

Heine et al. (1979: 88, endnote 29) report that Ehret suggests that Proto Kalenjin *raːro 'piece of hide used as container' is borrowed from Cushitic on the basis of the cognancy with Oromo *raro* 'skin of small animal'. Rottland (1982: 546) has Proto Kalenjin: *laːl 'leather bag'. The Oromo form is confirmed by Gragg (1982: 441) as *raroo* 'skin for sleeping on'. No cognate root was found in Konso, Mosiye, Dirayta. The forms may be cognate, but more evidence would be needed to propose transfer.

REEREE 'bat'

- PB *rɛɛrɛɛʕ 'bat' : PK *rɛːrɛːs : PS *raaraʕ

The PS form is solely based on Rendille *(raaráhhanyóy)*. The origin of the Rendille word is unknown. While it is possible that Proto Kalenjin borrowed the item from a predecessor of Rendille, transfer in the opposite direction is phonetically more likely, with Proto Kalenjin *s* > Rendille *h* > *ħ* (cf. Lamberti 1986: 300 on *h > *ħ*). We see no evidence for transfer from Cushitic.

RIRIM 'kind of ant'

Heine et al. (1979: 88, endnote 29) reports that Ehret claims that this Proto Kalenjin root *ririm is related to Oromo *riirma* 'termite'. The root is not present in Ehret (1971), nor in Rottland (1982) for Proto Kalenjin. He has other roots for ants and termite, Proto Kalenjin/ Proto South Nilotic *taːR 'termite' (Rottland 1982: 557) and Common Datooga *malil* 'black ant' (Rottland 1982: 487), cf. Nandi *cɛː-lɪl-*

DOI 10.15460/auue.2025.98.1.368

yaːt ‘small black ant with curled, sometimes yellow, tail’ (Creider & Creider 2001: 57).
Gragg (1982) has *rimma* ‘ant, termite’ for Oromo and Konso has a related root *irm-att-a* (m) ‘termite’ (Black & Otto 1973: 48) and Dirayta has *írimaʃʃ* ‘termites; large family’ (Black 1973); Mosiye *ʔirimaʃa* ‘termite’ (Yibeltal 2018). The Proto Oromoid root is probably close to *irim-. Black (1974) does not reconstruct this root for Lowland East Cushitic.

The root is clearly Cushitic but we consider it too much of a stretch to relate a root like *irim- to the Nandi root element *lɪl*, or to Datooga *malil*, as it involves too many changes: addition of initial *r* through reduplication, deletion of final *m* in Kalenjin, metathesis in Datooga, adaptation of *r* as *l*.

ROOB ‘rain’

- PB *roob ‘rain’ : PK * roːp- : PS *roob

This item is widely attested in East Cushitic, with Proto East Cushitic *roob- ‘rain’ on the basis of Saho-Afar rob; Somali *roob*, Rendille *roob* ‘green country’, Boni *roob*; Oromo *roob-a*, Konso *roop-a*, Dirasha *roop(-a)* (Sasse 1979: 22), also Dullay Gawwada *ʔirraw*, Tsamakko *ʔerr-o*; Dhaasanac *ʔir* ‘rain’, Elmolo *ʔirri* ‘rain’. Rottland (1982: 310) reconstructs *rɔɔp for Proto South Nilotic; the root seems absent in East Nilotic.

Given the near-identical forms across the Cushitic languages, transfer from various East Cushitic sources to Proto South Nilotic is conceivable, Proto East Omo-Tana *roob, Proto Oromoid *roob, although less readily from Dullay or West Omo-Tana due to differences in the shape.

SAANI ‘lover’

- PB *saani ‘lover’ : PK *saːn : Somali *saani*

Rottland (1982: 546) reconstructs this item to Proto South Nilotic *saːn with the more specific meaning of accepted male marriage candidate. The semantics of the item lends itself to borrowing. Given the restricted attestation in Cushitic (only Somali), the direction of borrowing must have been from Kalenjin to Somali.

DOI 10.15460/auue.2025.98.1.368

SAAC ‘to butcher’

Heine et al. (1979: 88, endnote 32) report that Ehret proposes that Proto Kalenjin *saːc is borrowed from Yaaku-Dullay. This root does not feature in Rottland (1982) under ‘to butcher’. We have not found evidence for its reconstruction in Proto Kalenjin nor candidate cognates in Cushitic.

SAGAAL ‘nine’

- PB *sagaal ‘nine ’ ; PK *sakaːl : PS *saagal

Rottland (1982: 423, 548) reconstructs Proto Kalenjin: *sakaːl and Common Datooga: *ʃageːʃ* but does not reconstruct a Proto South Nilotic form because the correspondences are not regular. Black reconstructs Proto Lowland East Cushitic *sagal- before C ~ *sagl- before V. Highland East Cushitic has a different root, *honso (Hudson 1989: 105). Also, the Agaw languages have a different root, Proto Agaw *säʒ-ta built on the number ‘four’ (Appleyard 2006: 105); and the word in Beja is yet different again (Wedekind et al. 2021). The South Cushitic root *gwaleeli (Kießling & Mous 2003) is not obviously related.

Even though the root is restricted to Lowland East Cushitic within Cushitic, it is even more restricted in Nilotic. Most of East and West Nilotic and many other branches of Nilo Saharan construct 9 as 5 + 4. In East Nilotic, only Samburu has *sâːl*, probably from Dhaasanac *saal* or a similar Cushitic form. In Surmic, Suri and Mursi have *sakkal* and Kwegu and Me’en have *sa(a)l*, presumably also due to local contacts with Cushitic. According to Blažek (2018: 53) the East Nilotic (especially Maa) numerals of Cushitic origin were not transmitted via South Nilotic (as Heine et al. (1979) posit), but via Southeast Surmic, which is “more probable from the point of view of both phonetics and linguistic geography.”

The irregular differences between Kalenjin and Datooga are either due to two different transfer events or because of an irregular sound change in the numeral system. For both options it is difficult to build a scenario. There is no Cushitic language (group) that has/had *sh* in this root to accommodate for that initial consonant in Datooga. Moreover, there is no evidence for direct East Cushitic contact with Datooga. The Datooga numeral paradigm does not contain other forms that start in *sh* that could explain the *s* > *sh* change as paradigmatic levelling. Possibly the final *sh*, which is a Datooga reflex

DOI 10.15460/auue.2025.98.1.368

of final Proto South Nilotic *L influenced the initial *s* to become *sh* too as an instance of palatal harmony. Proto South Nilotic *L derives from Proto Nilotic *li or *ly (Dimmendaal 1988). Proto Kalenjin has *saka:l; its final *l* can derive from Proto Nilotic *L as well. While there is no evidence for final *i* in East Cushitic, Proto West Rift has *gwaleeli*. This could be the source of both the Datooga and the Proto Kalenjin forms, if we assume palatal harmony, subsequent metathesis, and loss of labialisation (which is attested in Teso-Turkana (Dimmendaal 1984: 71), and possibly played a role in South Nilotic too).

A bold suggestion would be to reconstruct Proto Cushitic **sagweeli* > **gwaseeli* > Proto West Rift **gwaleeli*, parallel to Proto Cushitic **sagweeli* > **sageeli* > Proto East Cushitic **sagal* and transfer of **sageeli* > *sageesh* (*sh* in Datooga < Proto South Nilotic before *-i*).[8]

The South Nilotic words for 'nine' originate in Lowland East Cushitic but the details of the transfer cannot be determined. The vowel length in the second syllable is shared in South Nilotic with Somali and Rendille only. This suggests East Omo Tana as a source.

SEKEM 'honey' 'bee'

Heine et al. (1979: endnote 32) mention this as a proposal by Ehret (p.c.) to be a transfer from Yaaku-Dullay, but Ehret (1970: 153; 1971: 116) claims a South Cushitic origin. No source is given but Ehret (1970: 153; 1971: 116) considers it a South Cushitic loan assuming an *-em* South Cushitic suffix; it is unclear which suffix he has in mind. Transfer from South Cushitic cannot be shown. The root is not mentioned in Ehret's (2008) article on Yaakuan.

Heine et al. (1979: endnote 32) link the root to Proto East Cushitic *zagm- 'honey'. Sasse (1979: 20) reconstructs this on the basis of Oromo *damm-a*, Konso *takm-a*, Gidole *tonk-(a)* (by metathesis); Gawwada, Harso, and Gollango *takm-o*; and Yaaku *sikáʔ*, pl. *sakm-ay*. There is no cognate in East Omo-Tana. The words for 'bee' in Dullay are different: Gawwada *t'oonaq-o* and Tsamakko *ts'onaq'o*. Rottland (1982: 426, 532) reconstructs this root for Proto Kalenjin as *sɛkɛm

8 Bianjida Datooga has *shigeeshi* with a final *i* but it would entail transfer of a form (**sageeli*) that is too old for a Proto West Rift source; alternatively, it could have been transferred from East Omo-Tana. The problem with the latter is that there is no (other) evidence for contact between East Omo-Tana and Common Datooga so far. Another unexplained *sh* in Datooga may be *ɨŋaʃ* 'python', cf. POD *ɪŋaL (Rottland 1982: 458) – lower case *l* in *ŋal* in the index on p. 550 must be a typo.

DOI 10.15460/auue.2025.98.1.368

'bee'. It is hard to imagine how one can mistake honey for a bee and borrow a word for 'bee' in the meaning of 'honey'. Yet that is what seems to have happened here. A scenario how this can happen was provided to us by Sara Petrollino (p.c. October 6, 2023). She noticed that the Arbore people, while speaking Hamar, did exactly this, using the Hamar word 'bee' for 'honey'. Some Hamar speakers explained it to her as that these people were speaking *angúde aafó*, the label for an in-group speech register veiling messages specifically used by high status Hamar elders. While in a similar fashion, initiated boys during their seclusion in the wilderness would visit homes and demand milk, meat and honey using *maaz aafó*, a register in which they invert meaning and ask: "Don't give me bees" to mean "Give me honey". Moreover, in Datooga, the plural of 'bee' is polysemous in that it could also mean 'honey'.

Proto Kalenjin for 'honey' is different, *kʊʊm (Rottland 1982: 354) which cannot be shown to be inherited from Proto South Nilotic as Datooga has *mal* for 'honey', and that as well is borrowed from a Cushitic source, probably West Omo-Tana, see MALAB. Tosco & Sands (2022) suggest that this root is transferred into Proto-Yaaku-Dullay from a source genetically related to Hadza which has *sok'omo* 'to eat honey' and they propose that an earlier Hadza form would have been close to *sakx'omo with subsequent vowel harmony and merger of *kx'* and *k'* in the development of Hadza and on the Cushitic side the common Cushitic rule of vowel reduction in a syllable flanked by light syllables. They argue that this fits with the limited distribution within Cushitic (Yaaku-Dullay and Oromoid). Presence of this root in Oromoid would then be due to borrowing from Dullay.

Yaaku is a likely source provided the transfer is from the plural form for 'honey' into Proto Kalenjin because the Dullay languages have an initial *t* in the root. Alternatively, the *s* > *t* change occurred in Dullay after the transfer from Proto Yaaku-Dullay into Kalenjin due to late Konsoid influence, but this scenario is ruled out if we accept the proposal by Sands and Tosco for Pre-Hadza transfer into Proto Yaaku-Dullay. In conclusion, this is a good candidate for a transfer from Yaaku into Kalenjin. Note that quite a few of the Kalenjin languages have *k* as last consonant. Rottland (1982: 426) links it to words referring to 'area where wild honey can be collected'.

DOI 10.15460/auue.2025.98.1.368

SIZIET ‘eight’

- Heine et al. (1979): PB *siziet ‘eight’ : PK *sisiːt : PS *sizyeet

Rottland (1982: 431) reconstructs indeed *sisiːt for Proto Kalenjin, *sis(it) for Proto South Nilotic, and *sis for Common Datooga. And Heine (1978: 73) has *sízyéèt ‘eight’ for Proto East Omo-Tana, while Lamberti (1986: 239) posits *ʃizzet- (though without East Omo-Tana internal evidence for *ʃ* instead of *s*). If this is a transfer from Proto East Omo-Tana to Proto Kalenjin, this must have happened before *z* > *d* in Somali and *z* > *y* in Rendille, when Proto East Omo-Tana still had z. The root is also found in West Omo-Tana (Dhaasanac *síet* (Tosco 2001: 543), Arbore *suyé* (Hayward 1984: 417). Given the Proto South Nilotic reconstruction, Kalenjin inherited this word from Proto South Nilotic, and given the wider distribution in Cushitic the transfer was probably from Cushitic into Proto South Nilotic. The precise source in Cushitic was later than Proto East Cushitic, based on Proto East Cushitic *ʃiz[ħ]ent- ‘eight’ in Blažek (2018: 51). The item is not reconstructed to Proto East Cushitic (Sasse 1979) or Proto Cushitic (Ehret 1987), nor is any other root reconstructed for that concept; the same holds for Proto Nilotic in Dimmendaal (1988). This points to transfer from Proto Omo-Tana or Proto East Omo-Tana to Proto South Nilotic.

SOZZOM ‘thirty’

- PB sozzom ‘thirty’ : PK *sosom : PS: *sozzom*, Somali *soddon*

Rottland (1982: 424) reconstructs Proto Kalenjin *sasam. The reconstructed form **sozzom* in Heine et al. (1979) is probably projected on the basis of Somali *soddon;* it is not in Heine (1978).

The East Omo-Tana form is from an earlier contraction: *sízzàhh* ‘three’ + *tomman* ‘ten’. Rendille has a non-contracted and more recent compound in the reverse order *tomón séyahh*. This order of head-modifier is dominant in the Lowland East Cushitic languages and this order is apparent in all new compounds for these decimals, as in Konso *kunɗa sessaa* /ten three/ (Orkaydo 2013: 103), Dirayta, and Mosiye *hunɗa halɓata* /ten three/ (Yibeltal 2018). Within Oromoid, only Oromo kept the original contracted form, *soddoma* ‘thirty’. Most Lowland East Cushitic languages innovated with a transparent head-modifier construction of decimal followed by the digit:

DOI 10.15460/auue.2025.98.1.368

The Dullay languages[9], the West Omo-Tana languages Dhaasanac, Arbore and Elmolo innovated. Within Lowland East Cushitic, apart from Somali and Oromo, Bayso, Boni and Saho-Afar kept the old, contracted forms of digit-decimal while displaying a general head-modifier order within the noun phrase (Tosco 1994). Tosco (1994: 436) remarks on Proto East Cushitic nominal syntax that the numeral preceded the counted item which followed as a genitive, and Lowland East Cushitic languages mostly kept their postmodifying syntax, except that numeral was reanalysed in all languages except Somali as a typical modifier and was therefore shifted after its head.

Hence, these contracted forms on the basis of the modifier-head order are older and in common with Highland East Cushitic where the dominant noun phrase order is modifier head. Burji innovated the word for 'three', *fadi*; its expression for thirty is *fadiy-ttan*, the same construction as in the rest of Highland East Cushitic (Hudson 1989). Sasse (1982: 69) reports that the Burji form *fadii-ttánna* 'thirty' /3-10/ from *fadiya* 'three' which he reports is probably connected to Proto East Cushitic *sazħ 'three' but warning that *s* > *f* in Burji is not regular. This is despite the fact that Proto East Cushitic had head modifier order (Tosco 1994: 436). The same is valid for the Agaw languages: Proto Agaw *säɣwä-/səɣwä- 'three' + *-täŋən, 'tens' (Appleyard 2006: 135, 60). In Awngi, the new form for ten, *cká*, has replaced the old one in the expression of the decimals, hence *ʃúɣácká* 'thirty', while Conti-Rossini (1904) still reports forms with the old ending *-reŋ* for tens (Appleyard 2006: 135, 72).

The numeral cannot be reconstructed up to Proto South Nilotic level. Datooga has a construction 20 + 10 for thirty, *digdam aa daman* (Griscom ed. 2018: 17). The Kalenjin languages have an unexplained vowel length in V1 in three varieties (Rottland 1982: 424), and must have borrowed the numeral before Somali *zz* > *dd* and *m* > *n*.

Proto Kalenjin borrowed the word for thirty from a Cushitic source. This source is possibly Proto East Omo-Tana. A higher level cannot be excluded but is difficult to argue for because many Lowland East Cushitic languages innovated a construction for the decimals. Moreover, whether the transfer was to Proto Kalenjin, or to Proto South Nilotic is difficult to determine because 'thirty' is constructed as

9 Tsamakko, like Datooga, has a transparent construction 20 + 10 for 'thirty'.

DOI 10.15460/auue.2025.98.1.368

20 + 10 in Datooga. If this is a later Datooga innovation, the Proto South Nilotic form would have been lost.

SUBEEN 'ewe'

- PB *subeen 'virgin ewe' : PK *supeːn : PS *saben, Rendille *subén*

Heine (1978: 94) reconstructs this item to Proto East Omo-Tana on the basis of Somali and Rendille; Lamberti (1986) does not contain it. The reconstruction of the first vowel as *a* is based on Somali. Most attestations of this root have *u*, in Cushitic and in Nilotic and Bantu. The item is also found in Maasai *e-súpèn* 'ewe' (Payne & Kotikash 2008), Samburu (East Nilotic) *súpen*, with somewhat wider semantics such as 'young female sheep, goat, or donkey'. It is also found in Gawwada: *supun-t-e* 'young female sheep', Yaaku *supin* 'female goat which has not yet given birth' and a similar form is Konso *sukeen-ta* 'female sheep which hasn't given birth' belongs here too, though the velar place of articulation of the second consonant in Konso is unexplained so far, if it is cognate or related at all. This item is also widely attested in Northeast Bantu.

If the transfer into Proto Kalenjin was from Cushitic, the first vowel *(u)* suggests Rendille as the source. However, the reconstruction of the vowel *a* in Proto East Omo-Tana might have to be reconsidered. Sosal (in prep.) reconstructs *sVben for Proto East Omo-Tana in which the inserted vowel V can be either [a] (as it is in Somali) or [u] and assimilated to the labial consonant. Also, the various Kalenjin varieties pose problems concerning the vowel of the initial syllable. Various lects (Sapiny, Kony, Bong'om, Pok and Päkot) have *i* instead of *u* Rottland (1982: 435) possibly due to some front vowel harmony, or under influence of Turkana *sipaanit* 'drove, flock, herd' (Barrett 1990), whether cognate or not. The most likely scenario is a transfer from Pre-Rendille into Proto Kalenjin. The vowel *a* in Somali is problematic in this scenario. An alternative scenario or pattern of transfers would be from Rendille to Samburu to Maasai to Chaga, from Rendille to Kalenjin; from Maasai to Yaaku. This would leave open how Gawwada and Konso got the root with *u* as the first vowel. Alternatively, this root travelled from Maasai to Rendille, Samburu, Chaga, Kalenjin, and Yaaku.

DOI 10.15460/auue.2025.98.1.368

TEEKEEU ‘elbow’

Heine et al. (1979: 88 endnote 32) give this as Ehret’s proposal for Proto Kalenjin *teːkeːu being a loan from Yaaku-Dullay. We do not accept this to be a Cushitic loan in South Nilotic. Rottland (1982: 535) has Proto South Nilotic and Proto Kalenjin *ku(ː)tuŋ. The more solid and recent reconstruction by Rottland (1982) is different from the one proposed by Ehret and quoted in Heine et al. (1979) and can only be related to the Yaaku word *t’ekɛ* ‘arm’ (Baader nd.) if metathesis applied. Metathesis is unlikely to occur in a situation of transfer and must then be assumed to be a later development in the receiving language, but this again is unlikely given the fact the root is reconstructed at Proto South Nilotic level.

The Dullay words for elbow, Gawwada *t’ihil-e* and Tsamakko *ts’ekile*, are to be linked to Yaaku *t’okl-e?* ‘forearm’. There is an additional competing root for ‘armpit’ which is likely to be sound symbolically related to tickling: Gawwada *toloqoloq-itt-e*, Ma’a/Mbugu *ki-dóghólosá* ‘armpit’, and Hadza: *tl’ok’oto-ko*.

TIZZABA ‘seven’

- PB *tizzaba ‘seven’ : PK *tisAp : PS *tVzzoba

Rottland (1982: 443) reconstructs indeed *tɪsɑp for Proto Kalenjin and *isub for Common Datooga. The differences are irregular and therefore he does not reconstruct at Proto South Nilotic level, but it is evident that the reconstructed words in these branches are related. The Kalenjin languages show different ATR values of *i* which could be indicative for later loans, as is the difference *a/u*.

Transfer must again be at least early at Proto East Omo-Tana level before *zz* developed into *dd* in Somali or *y* in Rendille. The vowel of the first syllable is *i* in South Nilotic, while it is *o* in most Cushitic sources; the *e* in Rendille is presumably due to the *y* from *zz*. Dhaasanac and Elmolo have *i* but have lost the *z*; the development within West Omo-Tana is unclear; Arbore has *u*, *tuzba*, while Elmolo *tiipa* and Dhaasanac *tiiya* have *i*; West Omo-Tana is an equally plausible and problematic source. Yaaku *tisibo?* is the closest in form to the Datooga shape. It is impossible, however, to know what the shape of this word was in Proto Dullay/Yaaku because the Dullay languages innovated: Gawwada *tahan*; Tsamai *taħħan*. The root is not found within Cushitic outside of Lowland East Cushitic. Highland

DOI 10.15460/auue.2025.98.1.368

East Cushitic has *lamala (Hudson 1989: 131); South Cushitic has *faanqu, Proto Agaw has PA *läŋä-tta/laŋä-tta- formed by means of the suffix *-tta on the base of *läŋä/laŋä 'two' (Appleyard 2006: 120). Beja has *asarama* (Wedekind et al. 2021: 20).

Nilotic and Nilo-Saharan numerals between 6 and 10 tend to be based on 5; the South Nilotic languages have innovated in this respect and taken the number seven from a Lowland East Cushitic source. It is difficult to determine which one and to explain the deviant form in Datooga. The loss of the initial t in Datooga cannot be explained by a transfer from a different source since all Lowland East Cushitic cognates have an initial *t*.

TOMON 'ten'

- PB *tamman 'ten' ; PK *tɑmɑn : PS *tomman

Rottland (1982: 438) reconstructs Proto Kalenjin/Proto South Nilotic *tɑmɑn; *dɑmɑn; and Common Datooga: *dɑmɑm. This is a transfer into Proto South Nilotic from East Omo-Tana or an earlier source, Proto Omo-Tana. Sasse (1982: 176) reconstructs *tom(m)an-/*tomn- 'ten' for Proto East Cushitic. Proto East Omo-Tana is reconstructed as *tomman with final stress by Heine (1978). Somali has undergone denasalisation of the middle consonant, *toban*; hence if the transfer is from Proto East Omo-Tana, this must have happened before that change. However, the West Omo-Tana forms are virtually identical: Arbore *tomon*, Dhaasanac *tommon*, Elmolo *tomon*. The transfer could easily have been from Proto Omo-Tana. The Oromoid languages have *kuɗan* for 'ten' (Oromo, Konso) but they had a *-tama~ *-toma form earlier as this is still discernible in the decimal compound numerals, as in Oromo *afur-tama* '40' and *dig-dama* '20', cf. AFARTAM, the first entry of this Appendix.

TOR 'spear'

- PB *tor 'spear': PK *tɔr 'to pierce, spear' : Rendille *tor* 'spear'

While not reconstructed to Proto East Omo-Tana (Lamberti 1986), this root is widely attested in East Cushitic: Somali: *teeri* 'spear' (Said 2013: 349), Konso: *tor-a* 'spear', Yaaku: *tɔɔr* 'spear', Elmolo: *tɔɔr* 'harpoon', Inner Mbugu *i-torú* 'spear' and also Amharic *t'or*.

Rottland (182: 447) reconstructs it with a long vowel, Proto Kalenjin: *tɔːr 'to pierce with spear'. As in the case of 'rain', this item is a noun in the presumed donor and a verb in Kalenjin. We suggest

DOI 10.15460/auue.2025.98.1.368

transfer from Rendille or Elmolo or Yaaku. If Yaaku was the source, then it must have been lost in Dullay (while Konso has it). Elmolo has specialised its meaning to 'harpoon' which is not surprising given their occupation as fishermen. Elmolo is an unlikely source for Proto Kalenjin on cultural social grounds (presently), but Proto West Omo-Tana must have had this item.

TUAALIO 'cow bell'

Heine et al. (1979: 88 endnote 29) present Proto Kalenjin *tua:lio 'cow bell' as suggested by Ehret to be cognate with Oromo *dauweli* 'bell'. Likewise, Rottland (1982: 450) has Proto Kalenjin *twa:l* 'Glocke eines Rindes'; this is different from the root *kor that is presented in the main text of Heine et al. (1979: 78) and has a separate entry in this list. The Oromo form is not yet confirmed (absent in Gragg 1982, Stroomer 2001) but Konso has *tawn-a (m)* 'a bell (used for boasting)' (Black & Otto 1973: 53); not attested in Dirayta, nor in Mosiye. The initial stop is phonetically voiced in Konso. We need more attestations to reconstruct a Proto Oromoid form. In the absence of those, a transfer from Proto Oromoid to Proto Kalenjin cannot yet be proposed. Transfer from East Omo-Tana is also possible: The item is also attested in Saho and Somali *dawan* 'bell', Sidamo *däwäla* and Bilin *dawalä*, as well as in Ethiosemitic (Ge'ez *dawal* 'bell', Tigre, Tigrinya, Amharic, Gurage *däwäl*; Ge'ez *dawwala* 'to ring a bell', Tigre *däwwäla, däwwäna*, Tigrinya, Gurage *däwwälä*, Harari *däwwäl* (Leslau 1991: 145). Leslau is of the opinion that there is not enough evidence to determine the direction of borrowing (Semitic to Cushitic or vice versa) or whether it is a common inheritance. Rendille *ḍáam* 'metal cattle bell' (Pillinger & Galboran 1999: 99) might be a loan from Somali (with *m* stemming from etymological nativisation of Somali word-final *n*, cf. Rendille *sam* 'nose', Proto East Cushitic *san (Sasse 1979: 24)). It is unlikely that Proto Kalenjin got the item from East Omo-Tana.

WAAR 'kid'

While Rottland (1982: 453) does not reconstruct this item to Proto Kalenjin, he lists it as a potential Proto Kalenjin-Omotik candidate with the meaning 'young goat' (1982: 464); it is attested in Omotik *war*, Akie *war*, Pokot *wáwà*, pl. *wárá* 'kid (in general)' (Crazzolara 1978: 360) is similar but not necessarily related. Proto East Omo-

DOI 10.15460/auue.2025.98.1.368

Tana *waħar ‘kid’ to Proto Kalenjin has been excluded in Heine et al. (1979) since the Kalenjin reflex does not show the otherwise regular sound change *a > *e/ɛ in pharyngeal environment (Heine et al. 1979: 88 endnote 41). However, we could imagine Proto East Omo-Tana *waħar < *waħr which would be perceived as *wɛhr* in Kalenjin ears, adapted to *wɛr* and due to assimilation realised as *wɔr* or *wɒr*.

(Y)EEL ‘male bovine’

Heine et al. (1979: 88, endnote 32) report on Proto Kalenjin *(y)e:l being from Yaaku-Dullay according to Ehret. The root is not discussed in Ehret (1971); Rottland (1982: 556) has *e:R ‘bull’ for Proto South Nilotic. No cognate was found in Gawwada, Tsamakko, nor Yaaku. The reported claim in Heine et al. (1979) cannot be verified nor supported.

ZIG ‘cow dung, mud’

- PB *zig ‘cow dung, mud’ : PK * sik : PS * zig

The motivation for PS *zig is unclear. Lamberti posits *d’iik’- > North Somali *dhiiqo* ‘mud’, Benadir *dhiiqo/dhiikho*, Garre *dheeqə*, Tunni *dhiigə*, from Proto East Omo-Tana *d’yook’- (Lamberti 1986: 348), cf. also Oromo: *dikee* ‘manure’ (the semantics of “dung” + “mud” combined in one lexeme is common in East Africa and also present in Tanzanian Cushitic, e.g., *tluumfuqáa (f) ‘mixture of mud and wet dung’). A shift from *d’* or *d* > *z* is however not attested in East Cushitic. Formally close to the Proto Kalenjin form and semantically not too distant is Elmolo *sɛk* ‘salty soil from the lake which animal feed on’ (Omondi & Otieno 2008: 16). This does not seem to be cognate with Somali *dhiiqo*, as *k’* > *k* does not seem to be attested (Sasse 1979: 54). As to the initial consonant, the fate of Proto East Cushitic *d’, *d'_1 in Elmolo is not yet known (Sasse 1979: 29). The Elmolo word looks promising as a source, but the opposite direction is equally possible. In sum, this item does not present evidence for a Cushitic origin of the South Nilotic form so far.

ʕIDD ‘claw, fingernail’

- PB *ʕidd ‘claw, fingernail’ : PK *sɪːɣ : PS *ʕiddi

Lamberti (1986: 261) mentions the item for Tunni, North Somali and May, but does not reconstruct it to Proto East Omo-Tana for unknown reasons. If it were reconstructed, the second consonant would have

DOI 10.15460/auue.2025.98.1.368

been *zz in Proto East Omo-Tana according to Lamberti (1986: 260). This item is problematic on two accounts: ʕ being borrowed as *s* is improbable, and the same holds for *dd* or *zz* being borrowed as *ɣ/R*. In our view the Cushitic and Proto Kalenjin items are not related.

DOI 10.15460/auue.2025.98.1.368

In memoriam Norbert Cyffer (1943–2025)

Der Tod Norbert Cyffers am 12. März 2025 kam gänzlich unerwartet für die afrikanistische Community und als ein verstörender Schock für seine persönlichen Freunde. Der Tod riss einen bis zum Ende produktiven Wissenschaftler aus einem erfolgreichen Leben, das, wie die zahlreichen Nachrufe im Internet zeigen, national und international durch fachliche Anerkennung und enge persönliche Beziehungen gekennzeichnet war.

Geboren wurde Norbert Cyffer am 16. Mai 1943 in Dortmund. Den „ruhrpöttischen" Dialekteinschlag in seiner Sprache konnte er trotz der langen Zeit, die er in Österreich gelebt hat, nie ganz verleugnen. Nach dem Abitur und einem Kurzausflug in die Soziologie an der Universität Kiel kam er 1965 nach Hamburg, um hier Afrikanistik zu studieren. Es ist vielleicht hilfreich, etwas näher auf diese Ära einzugehen, die sie die Afrikanisten ihrer Zeit nachhaltig geprägt hat, gerade weil die Afrikanistik damals ziemlich anders war, als sie es heute ist. Damals wurde in Hamburg das Vorlesungsgebäude mehrfach in einem Semester geräumt, weil Bombendrohungen von verirrten Linksgruppen eingegangen waren. Sit-ins und Go-ins waren an der Tagesordnung, Vorlesungen wurden gestört, kurz: die alte bundesdeutsche, auch universitäre, Ordnung wurde von studentischer Seite heftig in Frage gestellt. Erstaunlicherweise fand kaum etwas von alledem in der Hamburger Afrikanistik statt, die zu der Zeit das Musterbeispiel eines sogenannten Orchideenfachs war – unter anderem dadurch gekennzeichnet, dass man sich in einer fast studentenfreien Zone befand, in der das gleichzeitige Erscheinen zweier ernsthaft interessierter Studenten – Cyffer war zusammen mit Ekkehard Wolff im 11. Stock des sogenannten Philosophenturmes aufgetaucht – eine Sensation darstellte, die den Seminardirektor Prof. Lukas veranlasste, von den neuen Studenten als von „Castor und Pollux" zu sprechen.

Die kleine Zahl von Studenten (Studentinnen gab es noch nicht, das änderte sich erst, als Heide (Reboul-)Mirt, Veronika Six und Theda Schumann ihr Afrikanistik-Studium in Hamburg begannen) führte dazu, dass alle schon sehr früh in die Forschungsarbeit von Prof. Lukas mit einbezogen wurden. Da Lukas großen Wert darauf legte, dass alle seine Studenten nicht nur in der Zeit, in der er im Institut anwesend war, sich auch dort aufhielten, fanden sich diese

DOI 10.15460/auue.2025.98.1.431

automatisch in einer engen Gemeinschaft wieder, die sie auch über das Fachliche hinaus zusammenband, und in der jeder seinen eigenen Arbeitsraum im Seminar hatte. Lukas verteilte unter ihnen Aufgaben aus seinen verschiedenen, weit gestreuten Interessensgebieten. Cyffer fielen das Kanuri und die saharanischen Sprachen zu – eine Lukassche Entscheidung, die Cyffers weiteren Lebensweg bestimmt hat. Lukas übertrug ihm die Bearbeitung der Syntax dieser Sprache, eines Bereiches, dem er in seinem eigenen grundlegenden Buch „A study of the Kanuri language“ (1937) ganze 16 Seiten gewidmet hatte. Und Cyffer lieferte: Seine Dissertation trug den Titel „Syntax des Kanuri“ (1973, publiziert 1974). Er hatte für sie umfangreiche Feldforschung in Maiduguri durchführen können, was für Hamburger Verhältnisse ungewöhnlich war. Bis dahin musste sich der Doktorand erst seine Sporen mit seiner Dissertation verdienen, bevor er „hinaus“ ins Feld reisen durfte.

Mit der fertigen Dissertation und seinen Afrika-Erfahrungen im Hintergrund konnte Cyffer seine Forschungen am Kanuri 1974 bis 1981 im Sprachgebiet des Kanuri fortsetzen. Er bekam eine Stelle an der Ahmadu Bello Universität in Zaria, an deren Dependance in Kano und dann an deren Zweig in Maiduguri, bevor auch Maiduguri seine eigene Universtät bekam. Cyffers Haus in Maiduguri wurde zum logistischen Stütz- und sozialen Mittelpunkt aller Hamburger Afrikanisten, die ihren Forschungen in Nordnigeria nachgingen, nicht zu vergessen die Abende im Maiduguri Social Club. Jahrelang haben wir über eine Notiz am dortigen Schwarzen Brett amüsiert: „Committee lost, sentimental value only“.

Cyffer arbeitete in Maiduguri zusammen mit John Hutchison an der Standardisierung der Orthographie des Kanuri, einem Projekt, das die Zusammenarbeit mit den muttersprachlichen Autoritäten dieser Sprache zwingend voraussetzte und ein großes Maß an Takt und Einfühlungsvermögen erforderte. Wie sehr er dabei Respekt und Anerkennung der traditionellen Elite gewann, zeigt sich darin, dass er den traditionellen Ehrentitel „Shettima Luggama Kanuribe“[1] durch den Shehu von Borno (2005) verliehen bekam, die erste derartige Verleihung an einen Europäer. Die Schaffung der Kanuri-Orthographie und deren Einführung in den Kanuri Grundschulunterricht ist ein dauerndes Verdienst, das durch ein

1 Cyffer & Hutchison (1990: 161): “*Shettíma Kanúribe* Honorary title given to a learned *malam* or outstanding people.”

DOI 10.15460/auue.2025.98.1.431

Kanuri-Wörterbuch (1990) und eine große Grammatik (2023) an Bedeutung und Nachhaltigkeit gewann. Zusätzlich arbeitete Cyffer in deutscher Afrikanistik-Tradition sprachvergleichend am Kanuri. Innerhalb der saharanischen Sprachfamilie erregten besonders die verschiedenen Verbalklassen seine Aufmerksamkeit, weil deren Zahl in den einzelnen Idiomen differierte, und es möglich schien, nachzuvollziehen, wie es zu dieser Situation gekommen war. Cyffer erarbeitete in Verfolgung der Arbeiten von Lukas ein Modell, das zeigte, wie die Entwicklung in den einzelnen Sprachen verlaufen sein könnte.

Parallel zu den rein sprachwissenschaftlichen Fragen beschäftigte Cyffer sich, ausgehend von der Sprachsituation in Nigeria, intensiv mit Fragen der Sprachsoziologie und Sprachpolitik, einem Feld, das er als zu wichtig erachtete, um es allein der Politik zu überlassen. Immer wieder betonte er die Möglichkeiten, die sich in der Verwendung der Muttersprachen in der allgemeinen Bildung boten. Er vermochte auf diese Weise, in seiner Arbeit Sprachwissenschaft und die Anwendung ihrer Ergebnisse in der Praxis in beispielhafter Weise zu vereinen.

Nach sieben Jahren Forschungs- und Lehrtätigkeit in Nigeria erhielt Cyffer 1982 den Ruf auf die Professur für Afrikanische Philologie an der Universität Mainz und 1994 den Ruf auf das Ordinariat für afrikanistische Sprachwissenschaft an der Universität Wien, das er bis zu seiner Emeritierung mit viel Engagement innehatte. Im Rahmen dieser Tätigkeiten lag ihm die akademische Lehre besonders am Herzen. Die zahlreichen Reaktionen auf die Nachricht von Cyffers Tode, gerade von ehemaligen Schülern aus dem Kanuri-Sprachraum, zeigen, wie erfolgreich seine Tätigkeit auf diesem Gebiet war.

Die früheren Hamburger Afrikanistik-Kommilitonen schlugen nach Abschluss ihrer Studienzeit sowohl fachlich wie auch räumlich getrennte Lebenswege ein. Dennoch erwies sich diese Zeit als ein Band, das sie über die Jahrzehnte zusammen gehalten hat. Immer wieder traf man sich bei Konferenzen, Afrikanistentagen, Kolloquien und Workshops und konnte dort die alten Verbindungen wieder aufnehmen und fortführen. Von nun an wird Cyffer bei diesen Gelegenheiten fehlen. Wir werden ihn sehr vermissen.

Unser Mitgefühl gilt seiner Frau Marita.

Ludwig Gerhardt

DOI 10.15460/auue.2025.98.1.431

Bibliographie

Cyffer, Norbert. 1974. *Syntax des Kanuri: Dialekt von Yerwa (Maiduguri)*. Hamburg: Buske.

Cyffer, Norbert & John Hutchison (eds). 1990. *Dictionary of the Kanuri Language.* Dordrecht & Maiduguri: Foris.

Cyffer, Norbert & Shettima Umara Bulakarima. 2023. *Reference grammar of Kanuri (Nigeria)*. Köln: Rüdiger Köppe Verlag.

Lukas, Johannes. 1937. *A study of the Kanuri language: Grammar and vocabulary*. London: Oxford University Press.

DOI 10.15460/auue.2025.98.1.431