WOPAL 2

WORKING PAPERS IN AFRICAN LINGUISTICS

Working Papers in African Linguistics vol. 2

Working Papers in African Linguistics (WoPAL) vol. 2

This work is part of the research output from the joint research projects “A new perspective on descriptive linguistics in Africa based on the translingual ecology” (2021-2024), and JSPS’s Core-to-Core Program: B. Asia-Africa Science Platforms titled “Establishment of a Research Network for Exploring the Linguistic Diversity and Linguistic Dynamism in Africa” (2018-2021). The publication of this volume is financially supported by the JSPS Grant-in-Aid for Scientific Research (B) (23K25319) and “Comprehensive Research in Language based on Descriptive and Philological Methods” hosted in Center for Language Studies, Kyoto Sangyo University (K2001).

この出版は，アジア・アフリカ言語文化研究所共同利用・共同研究課題「多言語混在状況を前提としたアフリカ記述言語学研究の新展開 (jrp000268)」(2021-2024)，および JSPS 研究拠点形成事業 (B.アジア・アフリカ基盤形成型)「アフリカにおける言語多様性とダイナミズムに迫るアフリカ諸語研究ネットワークの構築」(2018-2021) の成果にもとづくものです．また，科学研究費補助金 (基盤 B)「パラメター連動に基づくバントゥ諸語類型論：多様性と普遍性の原理的理解に向けて」(23K25319)，および京都産業大学総合学術研究所ことばの科学研究センター経費「記述と文献に基づくことばの総合的研究」(K2001) の援助を受けています．

Research Institute for Languages and Cultures of Asia and Africa (ILCAA),
Tokyo University of Foreign Studies
3-11-1, Asahi-cho, Fuchu-shi, Tokyo, 183-8534 Japan

KYOTO SANGYO UNIVERSITY
Center for Language Studies, Kyoto Sangyo University

Printed in Japan by Nihon Root Printing & Publishing Co., Ltd
ISBN 978-4-86337-545-1

Table of Contents

Palatalization and velarization in Rukiga pronunciation

Shigeki Kaji
Kyoto Sangyo University

Abstract
This paper offers a detailed description of palatalization and velarization in Rukiga (or orúkiga), a Bantu language of southwestern Uganda. Palatalization of consonants by the semivowel **y** [j] and velarization of consonants by the semivowel **w** [w] are characteristic of Rukiga pronunciation, but they are veiled under the common orthography Runyankore-Rukiga, which heavily weights Runyankore, a northern neighboring language closely related and similar to Rukiga but one that does not undergo consonants' palatalization and velarization by semivowels. The paper also reveals that Rukiga has much palatalization and velarization in common with Kinyarwanda, a southern neighboring language.

Keywords: Rukiga, Runyankore, Bantu, palatalization, velarization

1. Introduction

1.1. A puzzle

The author encountered the word **mákunyabiri** [mákuɲaβiɾi] "twenty" when describing a language called Chitembo in 1976 in the eastern part of the Congo (then Zaire).[1] No doubt this word comes from **mákumí** [mákumí] "tens" followed by the enumerative adjective **abiri** [aβiɾi] "two" as in example (1). Curiously enough, however, it was not the expected **mákumyabiri** [mákumjaβiɾi] but **mákunyabiri** [mákuɲaβiɾi], that is, with **mya** [mja] changed to **nya** [ɲa]. This word puzzled the author.

(1) a. mákunyabiri [mákuɲaβiɾi] "twenty"
b. mákumí[2] abiri
c. tens6[3] two6

[1] Chitembo is a Bantu language. From the residents' ledger at that time, the author estimated its speakers at 40,000 to 50,000. However, Lewis *et al*. (eds.) (2015) give 150,000 as the number of speakers.
[2] The high tone of **mí** of **mákumí** becomes low by assimilation to the low tone of **a** of **abiri**.
[3] Chitembo is a class language. The number after nouns is the specific noun's class number. Other word categories (e.g., adjectives, verbs) also take a prefix according to the noun with which they agree grammatically. The same rule applies to Rukiga, also a Bantu class language. The prefix **chi-** of the word **Chitembo** is the class 7 noun prefix that indicates the language. The same is done with the class 11 prefix **ru-** in Rukiga and Runyankore.

In 2017, 41 years later, the author encountered the same phenomenon when studying Rukiga, a Bantu language of southwestern Uganda. In Rukiga too, **mya** [mja] is changed to **nya** [ɲa], as in (2).

(2) a. makúmya:biri [makúɲa:βiɾi] "twenty"
b. makúmi abiri
c. tens6 two6

In Chitembo, **mákunyabiri** "twenty" was this phenomenon's sole example. In Rukiga, in contrast, examples abounded, and faced with them, the author came to understand that this phenomenon is explained as palatalization of **m** by the semivowel **y** [j]. Therefore, this paper aims to provide a detailed description and analysis of palatalization and related velarization in Rukiga pronunciation.

1.2. The Runyankore-Rukiga orthography

Although the paper deals with Rukiga, the relationship with its neighboring language Runyankore, spoken north of it, is another motivation. Rukiga and Runyankore[4] are so closely related that they usually carry the label Runyankore-Rukiga, and they share a common orthography, also termed Runyankore-Rukiga. Importantly, this orthography emphasizes Runyankore and hides Rukiga's pronunciation characteristics.

For example, the word for "head" is written in Runyankore-Rukiga orthography as **omutwe**, but its pronunciation differs between the two languages. In Runyankore, it is [omútwe] as written, while in Rukiga it is [omút^{k}we], with a slight [^{k}] inserted between [t] and [w]. This is Rukiga's feature, veiled under Runyankore-Rukiga orthography. An extreme example is the word for "dog," written as **embwa** in Runyankore-Rukiga orthography. Its pronunciation is [émbwa] in Runyankore but [émga] in Rukiga, a pronunciation that brings to mind a southern neighbor of Rukiga, Kinyarwanda, in which "dog" is pronounced [imgá].

Rukiga can be dually zoned as a northern dialect and a southern dialect. Roughly speaking, the northern dialect shares features with Runyankore, and the southern dialect with Kinyarwanda. This paper reflects Rukiga, especially its southern dialect, by illuminating its actual pronunciations, particularly palatalization and velarization.[5]

[4] According to Lewis *et al*. (eds.) (2015), Rukiga has 1,580,000 speakers, and Runyankore 2,330,000.

[5] Some of the several Runyankore-Rukiga dictionaries, for instance, Makerere University (2007), Museveni *et al*. (2012), and Mpairwe and Kahangi (2013), indicate pronunciations of segments and tone. However, no dictionary indicates Rukiga word pronunciation. For example, Makerere University's (2007) *Kashoboorozi y'Orunyankore-Rukiga* is rich in content and gives useful information about the language and culture of Runyankore-Rukiga. It is mindful of Rukiga words

1.3. Phonemes

This paper uses the Runyankore-Rukiga orthography in writing Rukiga words for ease of reference, except for vowel length marks and tone marking. Pronunciations are given in brackets when necessary. Example (3) lists Rukiga phonemes.[6]

(3) a. vowels: i, e, a, o, u
b. semivowels: y [j], w [w]
c. consonants: p [p], t [t], k [k]~[c], b [β]~[b], bb[7] [b], d [d], g [g]~[ɟ], f [f], s [s], sh [ʃ], h [h], v [v], z [z], j [ʒ], r [ɾ], (l [l],)[8] m [m],n [n] or [ŋ],ny [ɲ]
d. vowel length[9]: Long vowels are indicated by the vowel length mark ː.
e. tone: High tone and falling tone are indicated by an acute accent and a circumflex accent, respectively. Low tone is left unmarked.

2. Palatalization

The puzzling phenomenon mentioned in Section 1.1 is a type of Rukiga palatalization detailed in this section.

In the sequence **CyV**[10] in Rukiga, the semivowel **y** [j] gives rise to an alveolar consonant between **C** and **y** under the effect of palatalization. That is, the semivowel **y** makes the tongue body move up and forward toward the hard palate to produce an alveolar consonant that is either **t** or **d** depending on the voice quality of **C**; the consonant is **t** when **C** is voiceless and **d** when **C** is voiced. If the semivowel **y** is not preceded by a consonant in the syllable, no alveolar consonant is produced, as in (4).

---

but still biased by Runyankore pronunciation, so the words "head" and "dog" are written as **omutwe** [omútwe] and **embwa** [émbwa].

[6] Runyankore distinguishes between **ky** [c] and **c** [ʧ], and the Runyankore-Rukiga orthography uses both **ky** and **c**. However, Rukiga does not have the sound [ʧ], and both **ky** and **c** represent [c], which may be written as **ky**.

[7] The appearance of /bb/ [b] is relatively rare, appearing mostly in borrowings from Swahili and English.

[8] Not an original phoneme of Rukiga, /l/ appears only when pronouncing foreign words, including those of Luganda.

[9] In addition to phonemically long vowels, Rukiga has phonetically long vowels. Short vowels become long by glide formation and compensatory lengthening and also by a nasal cluster that follows a vowel. In some conditions, phonetically long vowels are not realized as fully long, but half-long, especially when high tone does not fall on them, in which case the half-length mark ˑ is used. Phonetic lengthening of vowels does not occur at the end of words.

[10] C and V stand for consonant and vowel, respectively. Other abbreviations used are the following. Ank.: Runyankore, appl.: applicative, aug.: augment, caus.: causative, Clit.: clitic, Eng.: English, FV: final vowel, gen.pres.: general present, NPr: nominal prefix, n.pst.stat.: near past and resulting state, ObjPr: object prefix, pass.: passive, per.: person, perf.: perfective, pl.: plural, poss.adj.: possessive adjective, pres.: present, sg.: singular, subj.: subjunctive, SubPr: subject prefix, sub.rel.: subject relative.

(4) a. i-angye 9 *poss.adj.* > yangye "my"
b. o.mu.shá.ya[11] 3 "jaw"

From Section 2.1 to Section 2.20, we examine all the cases of **C** and **V** in **CyV** sequences. Succession of two semivowels like **CywV** is rare in Rukiga. We find only one case of **myw** sequence (see Section 3.20), and no **(C)wyV** sequence is found in our data.

In addition to palatalization, we must consider spirantization of consonants by **y**. The semivowel **y** can be from the causative suffix **-i-**,[12] which may cause spirantization, such that **...t-i-a** > **...tya** > **...sa**. However, spirantization does not always occur in **CyV** sequences. Description of palatalization must be done carefully with respect to spirantization.

Because they exhibit a complete set of examples, we begin with **byV** sequences, which involve the bilabial voiced consonant **b**; however, their voiceless **pyV** sequence counterparts lack some occurrences.

2.1. byV > b^{d}yV (> dyV)

By palatalization of **y**, **byV** sequences become **b^{d}yV**. Very often, the resulting **d** replaces **b**, which thereby disappears (**b^{d}yV** > **dyV**). Because Rukiga has five vowels, as indicated in (3a), all five possibilities of **V** of **byV** sequences are listed in (5) and some examples in (6).

(5) a. bya > b^{d}ya (> dya)
b. bye > b^{d}ye (> dye)
c. byi > b^{d}yi (> dyi)
d. byo > b^{d}yo (> dyo)
e. byu > b^{d}yu (> dyu)

---

[11] The dot indicates a syllable boundary.

[12] Rukiga has two causative suffixes, namely the short causative suffix **-i-** and the long causative suffix **-i:s-**/**-e:s-**. The latter follows vowel harmony; it is **-i:s-** when the preceding vowel is /a, i, u/ and **-e:s-** when the preceding vowel is /e, o/. Only the short form **-i-** pertains to our discussion of spirantization.

a. short form **-i-**
okwó:roba "to be soft, tender, easy"
okwó:robya *caus.* "to make soft, etc."
b. long form **-i:s-**
okutumba "to swell"
okutumbi:sa *caus.* "to cause to swell"
c. long form **-e:s-**
okushoma "to read, to study"
okushome:sa *caus.* "to cause to read, to teach"

(6)	a.	okubyâːma	[okub^{d}jâːma]	or	[okudjâːma]	"to lie down, to sleep"
	b.	ebyeshero 8	[eb^{d}jeˑʃeɾo]	or	[edjeˑʃeɾo]	"watering places for cows"
	c.	ebyîːbo 8	[eb^{d}jîːβo]	or	[edjîːβo]	"baskets to serve food"
	d.	ebyóːndo 8	[eb^{d}jóːndo]	or	[edjóːndo]	"mud (caused by rain)"
	e.	ebyúːfe 8	[eb^{d}júːfe]	or	[edjúːfe]	"species of fruit"

In **CyV** contexts, the causative suffix **-i-** does not result in spirantization but palatalization, of which examples are provided in (7). Notably, the **b** sound in **okúraba** [okúɾaβa] "to pass" and **okuziba** [okuziβa] "to be blocked" is a fricative [β]. But when it is palatalized by **y**, it becomes [b] (which the orthography would put **bb**); [b] is a stop consonant and strong in comparison to [β], which is fricative and weak. The resulting [d] is also a stop and therefore a strong consonant. In this way, palatalization is considered a fortition process of consonants.

(7)	a.	okúraba	[okúɾaβa]			"to pass"
		okúrabya *caus.*	[okúɾab^{d}ja]	or	[okúɾadja]	"to help to pass"
	b.	okuziba	[okuziβa]			"to be blocked"
		okuzibya *caus.*	[okuzib^{d}ja]	or	[okuzidja]	"to block, to choke up"

Even when **b** is preceded by a nasal, the same process as (5) can occur, in which case **b** is [b], and the nasal is [m] since it is homorganic to the bilabial **b**. In casual speech, **mbyV** sometimes becomes **n^{d}yV** as described in (8). That is, **b** drops, perhaps via becoming **m** by assimilation to the preceding **m**, the result being **mm**, which simplifies to **m**. Then, the nasal **m** becomes **n**, assimilating to the following **d** in place of articulation. Examples are given in (9). For some speakers, palatalization does not occur in **mbyV** sequences, and words are pronounced as written like **omukóroróːmbya** 3 [omukóɾoɾóːmbja] "rainbow," **émbyâːra** 9 [émbjâːɾa] "way of planting," etc.

(8) mbyV > mb^{d}yV > mm^{d}yV > m^{d}yV (> n^{d}yV)

(9)	a.	omukóroróːmbya 3	[omukóɾoɾóːmb^{d}ja]	(or [omukóɾoɾóːn^{d}ja])	"rainbow"
	b.	émbyâːra 9	[émb^{d}jâːɾa]	(or [én^{d}jâːɾa])	"way of planting"
		cf. *okubyâːra*	[okub^{d}jâːɾa]		"to plant"
	c.	okubûːmbya	[okuβûːmb^{d}ja]	(or [okuβûːn^{d}ja])	"to make mold"
		cf. *okubûːmba*	[okubûːmba]		"to mold"
	d.	okutembya	[okuteˑmb^{d}ja]	(or [okuteˑn^{d}ja])	"to make climb up"
		cf. *okutemba*	[okuteˑmba]		"to climb up"

2.2. bbyV

No **bbyV** sequence appears in our data.

2.3. pyV

No nominals with a **pyV** sequence are found in our data, but two verbs have **p** as their final radical consonant. If they were to take the short form **-i-** as their causative suffix, the structural description would be met, but their causative suffix is the long **-iːs-**/**-eːs-**; neither are any examples of **pyV** sequences found in verbs.

(10)	a.	okukopa < Eng. *copy*	[okukopa]	"to cheat"
		okukopeːsa *caus*. (*okukopya)	[okukopeːsa]	"to make cheat"
	b.	okunyâːmpa	[okuŋâːmpa]	"to break wind"
		okunyáːmpiːsa *caus*.(*okunyâːmpya)	[okuŋáːmpiːsa]	"to cause to break wind"

2.4. dyV > d^dyV (> dyV)

In Rukiga, **d** and **r** are originally the same, that is, **d** in postnasal position (as in **okusîːnda** "to get drunk") and **r** elsewhere. Today, we have two types of **dyV** sequence, one occurring in postnasal position, and the other deriving from **ryV** by palatalization (**ryV > r^dyV > dyV**). For the former type, we see no phonetic change, as in (11) and (12). The latter type is addressed in Section 2.16.

(11) n-d-i-a > ndya

(12)	a.	endyarya 9	[endjaˑdja]	"betrayer"
	b.	okúːndya	[okúːndja]	"to eat me"
	c.	éndya 9	[éndja]	"way of eating"
		cf. *okúrya*	[okúdja]	"to eat"

Although we see no phonetic change in **ndyV** sequences in which palatalization applies, when spirantization applies, **y** (<**i**) changes **dyV** to **zV**, with the process illustrated in (13) and examples given in (14). In this case, the long causative suffix is not used.

(13) d-i-a > dya > zya > za

(14)	a.	okusîːnda	[okusîːnda]	"to get drunk"
		okusîːnza *caus.*	[okusîːnza]	"to make sb drunk"
	b.	okufûːnda	[okufûːnda]	"to be tight, narrow"
		okufûːnza *caus*.	[okufûːnza]	"to make tight or narrow"

2.5. tyV

No **tyV** sequence occurs in our data. The causative suffix **-i-** causes spirantization, with the process in (15) and examples in (16).

(15) t-i-a > tya > sya > sa

(16)	a.	okúteta	[okúteta]	"to be brought up spoiled"
		okútesa *caus*.	[okútesa]	"to bring up a child spoiled"
	b.	okutogota	[okutogota]	"to make a boiling noise"
		okutogosa *caus*.	[okutogosa]	"to boil (potatoes, etc.) in water"
	c.	okutagata	[okutagata]	"to be warm"
		okutagasa *caus*.	[okutagasa]	"to heat water (for bathing)"

2.6. cyV

No example of **cyV** sequence is found in our data. Even so, one remark may be in order about the **c** sound. As stated in footnote 6, Runyankore distinguishes between **ky** [c] and **c** [ʧ] phonemically (see Kaji 2004). In Rukiga, although the spelling **c** rather than **ky** is usually used with the back vowels /a, o, u/, it is not an affricate [ʧ] as in Runyankore, but palatal [c], the same as **ky** [c].

(17)	a.	ekicô:nco 7	[ecicô:nco]	"gift"
	b.	omucu:cu 3	[omucu:cu]	"dust"
	c.	okúca[13]	[okúca]	"to break"
	d.	okucumba	[okucuˈmba]	"to give off smoke"

2.7. gyV

In Rukiga, the sequence **gi** is pronounced [ɟi]. That is, the consonant **g** itself is palatalized when followed by /i/ or /e/. See examples in (18). Therefore, nothing except normal palatalization of **g** happens when a **gi** sequence is followed by a vowel, where **i** becomes the semivowel **y**. We find one such case, (19), in which the consonant **d** is not produced between **g** and **y**. A **gyV** sequence can be found morpheme-internally as well. See (20).

(18)	a.	egi 9	[eɟi]	"this"
	b.	orwigi 11	[ogwiˈɟi]	"door"
	c.	wangye 1 *poss.adj*.	[waˈɲɟe]	"my"
	d.	igyema 9	[iɟema]	"trouser"

(19) erá:ngí émwe > erá:ngy'é:mwe [eɾá:ɲɟé:mŋe] "one color"

[13] This word is **okúcwa** in the northern dialect.

(20) a. egyambíya 9 < Sw. *jambia* [eɟaˈmbíja] "machete"
b. enjogyera 9 [enʒoɟera] "small bell attached to the neck of a dog"
c. egyoːra 9 < Sw. *jora* [eɟoːra] "packed clothes"

The causative suffix **-i-** results in spirantization as exemplified in (22), with the process in (21). Some verbs like (22c) have two endings, **-za** or **-zya**.

(21) a. g-i-a > gya > zya > za
(22) a. okuhaːga [okuhaːga] "to eat heartily"
okuhaːza *caus*. [okuhaːza] "to satisfy"
b. okurêːnga [okurêːŋga] "to disappear from sight"
okurêːnza *caus*. [okurêːnza] "to make go beyond"
c. okwôːga [okwôːga] "to to bathe (intr.), to swim"
okwôːza *caus*. [okwôːza] or okwôːzya [okwôːzja] "to wash clothes, etc."

2.8. kyV

In Rukiga, the sequence **ki** is pronounced [ci]. That is, **k** is palatalized when followed by /i/ or /e/, in the same way as its voiced counterpart **g** (Section 2.7), as exemplified in (23). Nothing except normal palatalization of **k** happens when a **ki** sequence is followed by a vowel; then **i** becomes the semivowel **y**. In (24), we can confirm that the **t** sound is not produced between **k** and **y**. A **kyV** sequence can be found morpheme-internally. See (25).

(23) a. ekintu 7 [eciˈntu] "thing"
b. entakiːka 9 < Sw. *dakika* [entaciːka] "minute"
c omukyéːka 3 [omucéːka] "mat of papyrus stem fibers"
d. ekitoːki 7 [ecitoːci] "banana"
(24) a. ki-angye 7 *poss.adj*. > kyangye [caˈŋɟe] "my"
b. eki-énda 7 > ekyêːnda [ecêːnda] "rag"
c. eki-oyá 7 > ekyóːya [ecóːja] "feather"
(25) a. orukyankya 11 [oɾucaˈɲca] "banana leaf"
b. ekikyâːnkya 7 [ecicâːɲca] "stomach of humans or animals"
c. ekyekyerérezi 7 [eceɾérezi] "light"

The causative suffix **-i-** results in spirantization as exemplified in (27a,b,c) with the process in (26a). In some cases, however, both spirantization and palatalization apply. See (27d,e), in which case the process is like that in (26b). The two causative forms in (27d,e) have the same meaning. The process (26b) works mostly in Runyankore-influenced

words.

(26) a. k-i-a > kya > sa
b. k-i-a > kya > sya > sᵗya

(27) a. okutánaka [okutánaka] "to vomit"
okutánasa *caus*. [okutánasa] "to make vomit"
b. okuhika [okuhika] "to arrive"
okuhisa *caus*. [okuhisa] "to deliver"
c. okwambuka [okwaˈmbuka] "to go across a road, river, etc."
okwambusa *caus*. [okwaˈmbusa] "to help to go across a road, river, etc."
d. okwôːnka [okwôːŋka] "to suckle (intr.)"
okwôːnsa *caus*. [okwôːnsa] or okwôːnsya [okwôːnsᵗja] "to suckle (tr.)"
e. okutabaːruka [okutabaːruka] "to escape from (danger)
okutabaːrusa *caus*. [okutabaːrusa] or okutabaːrusya [okutaβaːrusᵗja]
"to congratulate"

2.9. vyV

No example of **vyV** sequence is found in our data. The **v** sound is relatively rare in Rukiga.

2.10. fyV

No example of the **fyV** sequence is found in our data, but at first glance, one verb **okúfa** 'to die' appears to have **-fa** as its ending. However, its radical is **-fú-**, and because it takes the long form **-iːs-** as its causative suffix, the sequence **fyV** does not appear. See **okufiːsa** (< **o-ku-fú-iːs**-a) 'to cause to die.'

2.11. zyV > zᵗyV > (zV)

Our data did contain **zyV** sequences, but all of them derive from the sequence **g-i-V** (or **r-i-V**) by spirantization by the short causative suffix **-i-** (see Sections 2.7 and 2.16). Although causatives with the **-zya** ending are used, forms without **y**, that is **-za,** are more commonly used, at least in the southern dialect.

2.12. syV > sᵗyV > (sV)

The sequence **syV** gives rise to **sᵗyV** by palatalization of **y**, creating a slight [ᵗ] after **s**. However, variation occurs between **syV** and **sV**.[14] See (29a,b,c,d). The sequence **syV** produces **sᵗyV**. The sequence **sV** remains unchanged, that is, no palatalization process

[14] In some northern dialect speakers, no palatalization occurs in **syV**, and it is pronounced [sjV] without [ᵗ].

occurs. Most examples of this type are those with the final vowel **-o**. Our data lacks examples of the sequences **sye**, **syi**, and **syu**.

The tendency in Rukiga is to prefer **sV** to **syV**, with **syV** forms mainly used in the northern dialect. However, in two words, namely, for 'herd' (29f) and 'forehead' (29g), **syV** forms are commonly used: for **ísyo** [ísʲjo] "herd," perhaps because cattle are mostly raised by the Banyankore; for **obúsyo** [obúsʲjo] "forehead," the influence of orthography is conceivable. Concerning the word for "New Testament" in (29h), the **syV** form is normally used, but in its pronunciation, [ʲ] is not heard.

(28) a. sya > sʲya
b. sye > sʲye (no example)
c. syi > sʲyi (no example)
d. syo > sʲyo
e. syu > sʲyu (no example)

(29) a. omúso 3 [omúso] or omúsyo [omúsʲjo] "knife"
b. oruguso 11 [oɾuguso] or orugusyo [oɾugusʲjo] "broken piece of a clay pot"
c. okwosa [okwoˑsa] or okwosya [okwoˑsʲja] "to burn (tr.)"
d. okúsa [okúsa] or okúsya [okúsʲja] "to be cooked [of food]"
e. okutaba:rusa *caus*. [okutaβa:rusa] or okutaba:rusya [okutaβa:ɾusʲja] "to congratulate"
f. íso 5 [íso] or ísyo 5 [ísʲjo] "herd"
g. obúso14 [obúso] or obúsyo 14 [obúsʲjo] "forehead"
h. endaga:no énsya 9 > endaga:n'é:nsya [endaga:né:nsja] "New Testament"[15]

2.13. jyV

No example of the **jyV** sequence is found in our data.

2.14. shyV

No example of the **shyV** sequence is found in our data.

2.15. hyV

No example of the **hyV** sequence is found in our data. The causative suffix **-i-** spirantizes the preceding **h**, changing it to **s**, and the suffix vowel itself becomes the semivowel **y**. See (31). At this point there is variation; either the semivowel **y** disappears (30a), or it causes palatalization and creates a slight **t** after **s** (30b). Rukiga prefers the **sV** type, the **sʲyV** type being Runyankore-influenced.

[15] The pronunciation **endaga:no énsa** > **endaga:n'é:nsa** [endaga:né:nsa] is possible but rare.

(30) a. h-i-V > hyV > sV
b. h-i-V > hyV > sʲV

(31) a. okuruha [okuɾuha] "to get tired"
okurusa *caus*. [okuɾusa] or okurusya [okuɾusʲja] "to make sb tired"
b. okuta:ha [okuta:ha] "to go in"
okuta:sa *caus*. [okuta:sa] or okuta:sya [okuta:sʲja] "to take in"

2.16. ryV > r^{d}yV > dyV

The sequence **ryV** becomes [djV], with **r** replaced by **d**. That is, the **r** sound, which is a liquid and weak, becomes the strong stop [d] by palatalization of **y**.

(32) a. rya > dya
b. rye > dye
c. ryi > dyi
d. ryo > dyo
e. ryu > dyu

(33) a. ryangye 5 *poss.adj*. [djaˈŋɻe] "my"
b. túrye *1st.per.sg.subj*. [túdje] "(that) we eat"
c. omúryi 1 [omúdji] "eater"
d. obúryo 14 [oβúdjo] "method"
e. ryu:míre 5 *n.pst.stat.* [dju:míɾe] "it is dry"

The causative suffix **-i-** changes **r** (originally ***d**) to **z** by spirantization. Examples are abundant. See (35). The spirantization process is indicated in (34).

(34) r-i-V > ryV > zyV > za

(35) a. okukórora [okukóɾoɾa] "to cough"
okukóroza *caus*. [okukóɾoza] "to cause to cough"
b. okúkira [okúciɾa] "to become cured"
okúkiza *caus*. [okúciza] "to cure, to heal"
c. okúkura [okúkoɾa] "to grow"
okúkuza *caus*. [okúkoza] "to bring up (a child)"
d. okusígara [okusígaɾa] "to remain"
okusígaza *caus*. [okusígaza] "to leave sth/sb"
e. okúbura [okúβuɾa] "to disappear"
okúbuza *caus*. [okúβuza] "to lose"

Interestingly, at this point, class 8 and class 5 markers can become homophonous; (36a) is a class 5 noun phrase, and (36b) is a class 8 noun phrase. The possessive adjectival

prefix of class 5 is **ri-**, and that of class 8 is **bi-**. However, when they are followed by a vowel (in this case, the possessive adjectival stem **-angye** 'my'), **ri-angye** becomes [djaˑɲɻe] and **bi-angye** also becomes [djaˑɲɻe] (see Section 2.1), thus becoming the same pronunciation.

(36) a. ibega ryangye [iβega djaˑɲɻe]
shoulder5 my5
"my shoulder"
b. ebitabo byangye [eβitaβo djaˑɲɻe]
books8 my8
"my books"

2.17. myV > m^{d}yV > n^{d}yV > n^{n}yV > nyV

The sequence **myV** becomes **nyV** [ɲV] in the last place. At the first stage, a slight **d** is produced under the palatalization effect of **y**. Then, the preceding nasal **m** assimilates to this alveolar **d**, which in turn becomes **n** assimilating to the resulting **n**. The resulting **nn** simplifies to **n**, which combines with the following **y**, resulting in **ny** [ɲ]. This process is observed particularly in class 4 nouns with the **mi-** prefix[16] when followed by a stem beginning with a vowel, as exemplified in (38).

(37) a. mya > nya
b. mye > nye
c. myi > nyi
d. myo > nyo
e. myu > nyu

(38) a. emi-áka 4 > emyâːka [eɲâːka] "years, harvest"
b. emi-énda 4 > emyêːnda [eɲêːnda] "clothes"
c. emi-íka 4 > emyîːka [eɲîːka] "smoke"
d. emi-oyo 4 > emyoyo [eɲoˑjo] "anuses"
e. emi-umá 4 > emyúːma [eɲúːma] "thirst"

The causative suffix **-i-** causes palatalization, not spirantization, in **myV** sequences. Examples are given in (39).

(39) a. okwêːma [okwêːma] "to be mounted [of cows, etc.]"
okwêːmya *caus*. [okwêːɲa] "to mount [of bulls, etc.]"

[16] The vowel element **e-** before the prefix **mi-** is called an augment, which has an article-like function.

b.	okúzima	[okúzima]	"to go out [of fire]"
	okúzimya *caus*.	[okúziɲa]	"to extinguish"
c.	okuguma	[okuguma]	"to be firm"
	okugumya *caus* .	[okuguɲa]	"to hold firmly"
d.	okwá:shama	[okwá:ʃama]	"to open wide [of the mouth, etc.]"
	okwá:shamya *caus*.	[okwá:ʃaɲa]	"to cause sb to open (the mouth)"

2.18. nyV

The **ny** [ɲ] sound is already palatal, and the sequence **nyV** is simply pronounced as [ɲV]. Examples are in (40). The causative suffix **-i-** palatalizes **n** to **ny** [ɲ], as in (41).

(40)	a.	okunya	[okuɲa]	"to defecate"
	b.	akanyegyero 12	[akaɲeɟero]	"uphill slope"
	c.	omushê:nyi 3	[omuʃê:ɲi]	"sand"
	d.	enyonyô:zi 9	[eɲoɲô:zi]	"star"
	e.	omû:nyu 3	[omû:ɲu]	"salt"
(41)	a.	okushúminkana	[okuʃúmiˈŋkana]	"to become tangled"
		okushúminkanya *caus*.	[okuʃúmiˈŋkaɲa]	"to tangle (tr.)"
	b.	okutâ:na	[okutâ:na]	"to have a breakup"
		okutâ:nya *caus*.	[okutâ:ɲa]	"to bust up a marriage"
	c.	okuhí:ngurana	[okuhí:ŋgurana]	"to go beyond"
		okuhí:nguranya *caus*.	[okuhí:ŋguraɲa]	"to make sb go beyond"

2.19. nyyV

No **nyyV** [ɲjV] sequence is found in our data. **Ny** is already palatal.

2.20. wyV

No **wyV** sequence is found in our data.

3. Velarization

In **CwV** sequences, the velar semivowel **w** [w] triggers velarization, so that the back of the tongue is raised toward the velum during articulation of **C** and gives rise to velar consonant **k** or **g** between **C** and **w**. The resulting consonant is **k** when **C** is voiceless and **g** when **C** is voiced. In some cases, **w** itself disappears. Velarization by **w** makes the preceding consonant strong in the same way as palatalization by **y**.

Velarization by **w** is noteworthy because **w** is a passive verb marker. Rukiga has two passive suffixes, namely the short form **-w-** and the long form **-ibw-**/**-ebw-**. The long form observes vowel harmony as **-ibw-** when the preceding vowel is /a, i, u/ and **-ebw-** when the preceding vowel is /e, o/. In most cases, the short form is used, but for monosyllabic

verbs (i.e., **-CV-** radical verbs), only the long form is used, as in (45c,d).

3.1. bwV > bgwV > bgV

(42) presents all five cases of **V** in **bwV** sequences in which **w** gives rise to **g** by velarization and **w** itself disappears. The resulting **g** totally replaces the original **w**. The examples given in (43) are nouns, and those in (44) are verb passives. The examples in (44a,b) are of the short passive suffix **-w-**, and those in (44c,d) are of the long passive suffix **-ibw-**/**-ebw-**. The **bwe** sequence can regularly be produced by putting passive forms in the subjunctive with the ending **-e**, as in (44b) and (44d). Notably, the fricative weak consonant [β] becomes a strong stop [b] by velarization of **w**, as exemplified in (44a,b).

(42) a. bwa > bga
b. bwe > bge
c. bwi > bgi
d. bwo > bgo
e. bwu > bgu

(43) a.	ekirí:bwa 7	[eciɾí:bga]	"sweet potato"
b.	obusabwe 14	[oβusabge]	"venom"
c.	obwí:na 14	[obgí:na]	"small holes, mesh"
d.	obwó:nko 14	[obgó:ŋko]	"brain"
e.	obwú:ma 14	[obgú:ma]	"dryness"
(44) a.	okure:ba	[okuɾe:βa]	"to see"
	okure:bwa *pass*.	[okuɾe:bga]	"to be seen (by sb)"
	aré:bwe *3rd.per.sing.pass.subj.*	[aɾé:bge]	"(that) s/he be seen (by sb)"
b.	okúroba	[okúɾoβa]	"to fish"
	okúrobwa *pass.*	[okúɾobga]	"to be fished (by sb)"
	eróbwe 9 *pass.subj.*	[eɾóbge]	"(that) it be fished (by sb)"
c.	okúrya	[okúdja]	"to eat"
	okurî:bwa (<o-ku-rí-ibw-a) *pass*.	[okuɾî:bga]	"to be eaten (by sb), to be edible"
d.	okúnywa	[okúŋwa]	"to drink"
	okunywê:bwa (<o-ku-nyó-ebw-a) *pass*.	[okuŋwê:bga]	"to be drunk (by sb), to be drinkable"
	ganywé:bwe 6 *pass.subj.*	[gaŋwé:bge]	"(that) it be drunk (by sb)"

When the resulting **bg** is preceded by a nasal (which is **m** since it is homorganic to **b**), **b** assimilates to the preceding nasal **m**, resulting in **mmgV**. Then, the geminate **mm** simplifies to a single nasal **m**. The process is shown in (45), and examples are given in

(46).

(45) mbwV > mbgwV > mbgV > mmgV > mgV

(46) a.	émbwa 9	[émga]	"dog"
b.	entumbwe 9	[entuˈmge]	"calf"
c.	okuyamba	[okujaˈmba]	"to help"
	okuyambwa *pass*.	[okujaˈmga]	"to be helped (by sb)"
d.	okusamba	[okusaˈmba]	"to kick"
	okusambwa *pass*.	[okusaˈmga]	"to be kicked (by sb)"

3.2. bbwV

No **bbwV** sequence appears in our data.

3.3. pwV > p^{k}wV > pkV

The sequence **pwV** becomes **p^{k}wV** by velarization of **w**. Then **w** drops, resulting in **pkV**. As the cause of velarization, **w** is totally replaced by **k**. Examples of the sequence **pwV** are few. All **pwV** sequences in (48) are preceded by the nasal **m**, which is the first-person singular object prefix meaning 'me'. More precisely, because of this **m**, **p** appears, which otherwise appears as **h**.[17] Borrowed from English, (49) is a sole example of the sequence **pwV** not preceded by a nasal.

(47) a.	pwa	> pka		
b.	pwe	> pke		
c.	pwi	> pki	(no example)	
d.	pwo	> pko	(no example)	
e.	pwu	> pku	(no example)	
(48) a.	zimpwerého 10 *n.pst.stat.*		[ziˈmpkeˈɾého]	"they finished to me"
	cf. *okúhwa*		[okúhwa]	"to finish (intr.)"
b.	orampwêːza?		[oɾaˈmpkêːza]	"can you see me clearly?"
	cf. *okuhwêːza*		[okuhwêːza]	"to see clearly"
(49)	okukopa < Eng. *copy*		[okukopa]	"to cheat"
	okukopwa *pass*.		[okukopka]	"to be cheated (by sb)"
	akópwe *3rd.per.sg.pass.subj.*		[akópke]	"(that) s/he be cheated (by sb)"

3.4. ndwV > nd^{g}wV (or ngwV)

Cases of **dwV** sequence are relatively rare because **d** normally appears after a nasal.

[17] More examples can be made with verbs whose stem begins with **h** if we use this **m** "me" at the same time.

Apart from passive forms in (52), only two noun examples (51a,b) in which **ndwV** becomes **ngwV** are attested. That is, **g** appears between **d** and **w**, and the resulting **g** replaces **d**. We know that the **d** of **endwâːra** 9 "sickness" (51a) is from **r** by its alternation with **oku-rwâːr-a** "to be sick," and noun examples are addressed in Section 3.16 along with **rwV** examples.

The **dwV** sequence of verb passives also undergoes velarization but only a partial process. Certainly **g** appears after **d** under the velarization effect of **w** but does not replace **d**, which therefore does not disappear.

(50) a. ndwa > nd^{g}wa (> ngwa)
b. ndwe > nd^{g}we (> ngwe)
c. ndwi > nd^{g}wi (no example)
d. ndwo > nd^{g}wo (no example)
e. ndwu > nd^{g}wu (no example)

(51) a. endwâːra 9 [eŋgwâːra] "sickness"
b. emáːndwa 9 [emáːŋgwa] "tools to communicate with spirits"

(52) a. okuhêːnda [okuhêːnda] "to break with hands"
okuhêːndwa *pass*. [okuhêːnd^{g}wa] "to be broken (by sb)"
b. okukûːnda [okukûːnda] "to love"
okukûːndwa *pass*. [okukûːnd^{g}wa] "to be loved (by sb)"
c. okuronda [okuɾoˑnda] "to look for"
okurondwa *pass*. [okuɾoˑnd^{g}wa] "to be looked for"
aróndwe *3rd.per.sg.pass.subj.* [aɾóːnd^{g}we] "(that) s/he be looked for (by sb)"

3.5. twV > t^{k}wV

In **twV** sequences, a slight [^{k}] appears between **t** and **w** by velarization of **w**. No further change occurs even when **twV** is preceded by a nasal, as in (54d).

(53) a. twa > t^{k}wa
b. twe > t^{k}we
c. twi > t^{k}wi
d. two > t^{k}wo
e. twu > t^{k}wu

(54) a. obutwa 14 [oβut^{k}wa] "poison"
b. omútwe 3 [omút^{k}we] "head"
c. okútwi 15 [okút^{k}wi] "ear"
d. entwíːga 9 [ent^{k}wíːga] "giraffe"

In addition to passive forms (see 55), **twV** sequences appear particularly in forms

involving the first-person plural prefix **tu-** "we, us" (see 56a,b) and the class 13 prefix **tu-** (see 56c,d) when they are followed by an element beginning with a vowel such as the tense marker **-a-**.

(55)	a.	okukúrata	[okukúɾata]	"to follow"
		okukúratwa *pass*.	[okukúɾatᵏwa]	"to be followed (by sb)"
	b.	okwêːta	[okwêːta]	"to call"
		okwêːtwa *pass*.	[okwêːtᵏwa]	"to be called (by sb)"
		ayéːtwe *3rd.per.sg.pass.subj*.	[ajéːtᵏwe]	"(that) s/he be called (by sb)"
(56)	a.	twarúha *1st.per.sg.pres.perf.*	[tᵏwaˈɾúhe]	"we have become tired"
	b.	twubíːra *1st.per.sg.gen.pres.*	[tᵏwuβíːɾa]	"we always dive"
	c.	twoːróbi 13 *n.pst.stat.*	[tᵏwoːɾóβi]	"they are easy, simple"
	d.	twangye 13 *poss.adj*.	[tᵏwaˈɲɟe]	"our"

3.6. cwV

The letter **c** represents the affricate [ʧ] in Runyankore, but is the palatal stop [c] in Rukiga. Although the northern dialect of Rukiga has verbs with the **-cwa** ending (e.g., **okúcwa** "to break, to tear") like Runyankore, the southern dialect on which this study is based prefers forms without **w** (e.g., **okúca** "to break, to tear" rather than **okúcwa**). Therefore, examples of **cwV** are few. In **cwV** sequences, [ᵏ] does not appear.

(57)	a.	cwa	>	cwa	
	b.	cwe	>	cwe	
	c.	cwi	>	cwi	(no example)
	d.	cwo	>	cwo	(no example)
	e.	cwu	>	cwu	(no example)

(58)	a.	amacwáːnsi 6	[amacwáːnsi]	"saliva, spit"
	b.	ekicwéːka 7	[ecicwéːka]	"half, portion"
	c.	encwéːra 9	[ɛɲcwéːɾa]	"cobra"
	d.	omucwêːzi 1	[omucwêːzi]	"mythical figure"

(59) okúca [okúca] or okúcwa [okúcwa] "to break, to tear"

3.7. gwV (> gᵍwV > gwV)

Sequences of **gwV** appear as if no change occurs in them. Possibly, however, velarization occurs, and the resulting geminated **gg** simplifies to a single **g**, that is, a change like the following: **gwV**> **gᵍwV**> **gwV**. This interpretation is suggested by its voiceless counterpart, the **kwV** sequence (Section 3.8), in which velarization may occur due to the velar **w,** although its effect is very weak. Examples of **gwV** sequence are in (61) and (62), and those in (62) involve verb passives.

(60) a. gwa > gwa
b. gwe > gwe
c. gwi > gwi
d. gwo > gwo
e. gwu > guː

(61)	a.	gwangye 3 *poss.adj.*	[gwaˈŋɻe]	“my”
	b.	engwe 9	[eŋgwe]	“leopard”
	c.	okugwira *appl.*	[okugwiˈra]	“to fall onto sth”
		cf. *okugwa*	[okugwa]	“to fall”
	d.	gworóbi 3 *pres.*	[gwoˈɾóβi]	“it is easy”
	e.	guːmíre 3 *n.pst.stat.*	[guːmíɾe]	“it is dry”
(62)	a.	okwéːsiga	[okwéːsiga]	“to trust”
		okwéːsigwa *pass.*	[okwéːsigwa]	“to be trusted (by sb)”
	b.	okusîːnga	[okusîːŋga]	“to win”
		okusîːngwa *pass.*	[okusîːŋgwa]	“to be defeated (by sb)”
	c.	okúsiga	[okúsiga]	“to leave behind”
		okúsigwa *pass.*	[okúsigwa]	“to be left behind (by sb)”
		asígwe *3rd.per.sg.pass.subj.*	[asígwe]	“(that) s/he be left behind (by sb)”

3.8. kwV (> k^kwV)

Most **kwV** sequences come from **ku-V** sequences, as in (64a,c,d,e), and velarization seems not to work in **kwV** sequences. However, if we listen attentively, we may perceive a slight [^{k}] in **kwi** sequences as [k^kwi] (see 64c), but mostly, nothing seems to happen. That is, even if [^{k}] may be produced by velarization of **w**, the resulting geminated [k^k] simplifies to a single [k]. Examples of **kwV** sequence are in (64) and (65), and those in (65) involve verb passives.

(63) a. kwa > kwa (k^kwa > kwa)
b. kwe > kwe (k^kwe > kwe)
c. kwi > kwi (k^kwi > kwi)
d. kwo > kwo (k^kwo > kwo)
e. kwu > kuː (k^kwu > kuː)

(64)	a.	ku-angye 15 *poss.adj.*	> kwangye	[kwaˈŋɻe]	“my”
	b.	omú-kwe 1	> omúkwe	[omúkwe]	“son-in-law”
	c.	oku-íba	> okwîːba	[okwîːβa] or [ok^{k}wîːβa]	“to steal”
	d.	oku-ónka	> okwôːnka	[okwôːŋka]	“to suckle”
	e.	oku-úma	> okûːma	[okûːma]	“to be dry”
(65)	a.	okuténdeka		[okuténdeka]	“to train”
		okuténdekwa *pass.*		[okuténdekwa]	“to be trained (by sb)”

b.	okutêːka	[okutêːka]	"to cook"
	okutêːkwa *pass*.	[okutêːkwa]	"to be cooked (by sb)"
c.	okureka	[okuɾeka]	"to leave, to abandon"
	okurekwa *pass*.	[okuɾekwa]	"to be left, abandoned (by sb)"
	arékwe 3*rd.pes.sg.pass.subj.*	[aɾékwe]	"(that) s/he be abandoned (by sb)"

3.9. vwV

No example of **vwV** sequence is found in our data. Neither is any verb with the **-va** ending, whose passive would have a **-vwa** ending.

3.10. fwV

No example of **fwV** sequence is found in our data. One verb with the **-fa** ending, namely **okúfa** "to die," actually has the radical **-fú-**, and because it is intransitive, it cannot be transformed to passive.

3.11. zwV > zᵍwV

Most **zwV** sequences come from passive forms in which the sequence **zwV** gives rise to **zᵍwV** by palatalization of **w**. Clearly, the word **ebisígarízwa** 8 [eβisígaɾízᵍwa] "residue of tea leaves" (67b) derives from a verb passive. The word **ebyóːgyézwa** 8 [ebᵈjóːɟézᵍwa] "wastewater" (67a) also seems to derive from a verb passive.

(66) a. zwa > zᵍwa
b. zwe > zᵍwe
c. zwi > zᵍwi (no example)
d. zwo > zᵍwo (no example)
e. zwu > zᵍuː (no example)

(67) a.	ebyóːgyézwa 8	[ebᵈjóːɟézᵍwa]	"wastewater"
b.	ebisígarízwa 8	[eβisígaɾízᵍwa]	"residue of tea leaves"
	cf. *okusígarizwa*	[okusígaɾizᵍwa]	"to be left for sb (by sb)"
(68) a.	okukyâːza	[okucâːza]	"to whet, to sharpen"
	okukyâːzwa *pass*.	[okucâːzᵍwa]	"to be whetted (by sb)"
b.	okúkiza	[okúciza]	"to heal"
	okúkizwa *pass*.	[okúcizᵍwa]	"to be healed (by sb)"
c.	okushémeza	[okuʃémeza]	"to operate"
	okushémezwa *pass*.	[okuʃémezᵍwa]	"to be operated (by sb)"
	ashemézwe *3rd.per.sg.pass.subj.*	[aʃemézᵍwe]	"(that) s/he be operated (by sb)"

3.12. swV > s^{k}wV

swV sequences become **s^{k}wV** by velarization of **w**, in which a slight [^{k}] appears after **s**. Although the sequence **swV** regularly appears in passives, as in (71), it is rare in other categories of words. No example of nominals has yet been found. In this regard, the English word *swamp* is noteworthy—often used because the Kigezi region has swampy areas. Interestingly, when Kiga people pronounce *swamp*, they insert the **k** sound in the same way as in their genuine Rukiga words. See examples in (70).

(69) a. swa > s^{k}wa
b. swe > s^{k}we
c. swi > s^{k}wi
d. swo > s^{k}wo (no example)
e. swu > s^{k}wu (no example)

(70) a.	swamp *Eng*.	[s^{k}wamp]	"swamp"
b.	sweet *Eng*.	[s^{k}wiːt]	"sweet"
c.	swear *Eng*.	[s^{k}wéə]	"swear"
d.	swallow *Eng*.	[s^{k}wɑ́lou]	"swallow"
(71) a.	okutánasa	[okutánasa]	"to make vomit"
	okutánaswa *pass*.	[okutánas^{k}wa]	"to be made vomit (by sb)"
b.	okwéːgyeːsa	[okwéːɟeːsa]	"to teach"
	okwéːgyeːswa *pass*.	[okwéːɟeːs^{k}wa]	"to be taught (by sb)"
c.	okuhíːmbiːsa	[okuhíːmbiːsa]	"to praise"
	okuhíːmbiːswa *pass*.	[okuhíːmbiːs^{k}wa]	"to be praised (by sb)"
	ahimbíːswe *3rd.per.sg.pass.subj.*	[ahiˈmbíːs^{k}we]	"(that) s/he bepraised (by sb)"

3.13. jwV > j^{g}wV

The **jwV** sequence becomes **j^{g}wV,** in which the velar semivowel **w** produces **g** between **j** and **w** by velarization. Passive forms are normally pronounced with this **g**, but examples are few. See (74).

(72) a. jwa > j^{g}wa
b. jwe > j^{g}we
c. jwi > j^{g}wi
d. jwo > j^{g}wo (no example)
e. jwu > j^{g}wu (no example)

(73) a.	omushójwa 3	[omuʃóʒ^{g}wa]	"species of spear grass"
b.	okwéːjweka	[okwéːʒ^{g}weka]	"to be able to dress oneself"
c.	enyaːrújwi 9	[eɲaːrúʒ^{g}wi]	"chameleon"

	d.	okujwa	[okuʒ^{g}wa]	"to bleed"
(74)	a.	okugunja	[okuguˈnʒa]	"to write lyrics"
		okugunjwa *pass*.	[okuguˈnʒ^{g}wa]	"to be written [of lyrics] (by sb)"
	b.	okubeːja	[okuβeːʒa]	"to carve"
		okubeːjwa *pass*.	[okuβeːʒ^{g}wa]	"to be carved (by sb)"
		gubéːjwe 3 *pass.subj.*	[guβéːʒ^{g}we]	"(that) it be carved (by sb)"

3.14. shwV > sh^{k}wV

The sequence **shwV** gives rise to **sh^{k}wV** by velarization of the velar **w**. Examples are relatively few, except for passive forms. See (77). Notably, even English loans like (76c) undergo this velarization.

(75)	a.	shwa	>	sh^{k}wa	
	b.	shwe	>	sh^{k}we	
	c.	shwi	>	sh^{k}wi	
	d.	shwo	>	sh^{k}wo	(no example)
	e.	shwu	>	sh^{k}wu	(no example)

(76)	a.	omúshwa 3	[omúʃ^{k}wa]	"termite"
	b.	múshushwe 9	[múʃuʃ^{k}we]	"shrew mouse"
	c.	eshwéːta 9 <Eng.	[eʃ^{k}wéːta]	"sweater"
	d.	eshwîːga 9	[eʃ^{k}wîːga]	"species of vegetable"
(77)	a.	okushiːsha	[okuʃiːʃa]	"to spoil"
		okushiːshwa *pass*.	[okuʃiːʃ^{k}wa]	"to be spoiled (by sb)"
	b.	okuhêːsha	[okuhêːʃa]	"to work iron"
		okuhêːshwa *pass*.	[okuhêːʃ^{k}wa]	"to be worked [of iron] (by sb)"
	c.	okúrasha	[okúraʃa]	"to fire, to shoot"
		okúrashwa *pass*.	[okúraʃ^{k}wa]	"to be shot (by sb)"
		aráshwe *3rd.per.sg.pass.subj.*	[aɾáʃ^{k}we]	"(that) s/he be shot (by sb)"

3.15. hwV

The sequence **hwV** does not produce any change and is pronounced without the effect of velarization. Normally, velarization occurs in **CwV** sequences, and seen in this light, the quality of **h** as a consonant is weak.

(78)	a.	hwa	>	hwa
	b.	hwe	>	hwe
	c.	hwi	>	hwi
	d.	hwo	>	hwo (no example)
	e.	hwu	>	hwu (no example)

(79)	a.	íhwa 5	[íhwa]	"thorn"
	b.	okuhwehura	[okuhweˈhuɾa]	"to whisper"
	c.	ogúhwire 3 s*ub.rel.*	[ogúhwiˈɾe]	"that which has finished"
(80)	a.	okuroha	[okuɾoha]	"to drown (tr.)"
		okurohwa *pass*.	[okuɾohwa]	"to be drowned (by sb)"
	b.	okura:ha	[okuɾa:ha]	"to castrate"
		okura:hwa *pass*.	[okuɾa:hwa]	"to be castrated (by sb)"
	c.	okúboha	[okúβoha]	"to bind"
		okúbohwa *pass*.	[okúβohwa]	"to be bound (by sb)"
		ebóhwe 9 *pass.subj.*	[eβóhwe]	"(that) it be bound (by sb)"

3.16. rwV > r^{g}wV > gwV

Sequences of **rwV** become **gwV** by velarization of **w** through the intermediate stage **r^{g}wV**. In **r^{g}wV** sequences, the **r** sound is expelled by the resulting **g**, which replaces **r**. If the **V** is **u**, the resulting **rwu** sequence is realized as **gu:**. The same velarization process applies to the word **endwâ:ra** 9 "sickness" listed in (51a), i.e., **ndwV** > **nd^{g}wV** > **ngwV**. That is, **g** appears after **d** even if preceded by the nasal **n** under the velarization influence of **w,** as in other examples. Then **d** drops, with the resulting **g** replacing it.[18]

Sequences of **rwV** cause intricate situations, especially when the class 11 prefix **ru-** is followed by a vowel. For example, (82a) **rwangye** [gwaˈɲɟe] "my" is the possessive adjective of class 11. Very curiously, however, the class 3 form of **gwangye** "my" is also [gwaˈɲɟe]. "My" of class 11 is written **rwangye** but pronounced [gwaˈɲɟe] in exactly the same way as the class 3 "my" [gwaˈɲɟe]. Worthy of note is that Kiga people write the class 11 adjective "my" as **rwangye** but pronounce it as [gwaˈɲɟe], never as [ɾwaˈɲɟe]. The name of the country Rwanda is spelled **Rwanda**, but its pronunciation is [gwaˈnda], not [ɾwaˈnda].

(81)	a.	rwa	>	gwa	
	b.	rwe	>	gwe	
	c.	rwi	>	gwi	
	d.	rwo	>	gwo	
	e.	rwu	>	gwu > gu:	

(82)	a.	rwangye 11 *poss.adj*.	[gwaˈɲɟe]	"my"
	b.	orwê:nda 11	[ogwê:nda]	"rag"
	c.	orwî:ri 11	[ogwî:ɾi]	"species of grass"
	d.	orwô:ya 11	[ogwô:ja]	"steam, vapor"

[18] This seems simpler than the following possible process: **ndwV** > **nd^{g}wV** > **nn^{g}wV** > **n^{g}wV** > **ngwV**. That is, **g** appears after **d** under the influence of **w** like the other examples, but **d** changes to **n** by nasal assimilation. Then the resulting **nn** sequence is simplified to a single **n**.

e. rwumíre 11 *n.pst.stat.* [guːmíɾe] "it is dry"

Because in (82), all examples are class 11 forms with the nominal prefix **ru-** (e.g., **orúkiga** "the Kiga language"), we can establish class 11 prefixes as **ru-** even though it is pronounced [gw] when a vowel follows. In passives too, we can discern the origin (i.e., underlying consonant) of [g], either **r** or **g**, because alternation occurs between [ɾ] and [g]. For example, in (83a), [g] of [okúʃagwa] "to be cut (by sb)" is determined as **r** in reference to **okúshara** [okúʃaɾa] "to cut" even though its pronunciation is [okúʃagwa]. Therefore, it is spelled with **r** as **okúsharwa** [okúʃagwa]. The same can be said of (83b) **okúkorwa** [okúkogwa] "to be done (by sb)" and (83c) **okutwâːrwa** [okut^{k}wâːgwa] "to be taken (by sb)."

(83)	a.	okúshara	[okúʃaɾa]	"to cut"
		okúsharwa *pass*.	[okúʃagwa]	"to be cut (by sb)"
	b.	okúkora	[okúkoɾa]	"to do"
		okúkorwa *pass*.	[okúkogwa]	"to be done (by sb)"
	c.	okutwâːra	[okut^{k}wâːɾa]	"to take"
		okutwâːrwa *pass*.	[okut^{k}wâːgwa]	"to be taken (by sb)"
	d.	zitwáːrwe 10 *pass.subj.*	[zit^{k}wáːgwe]	"(that) they be taken (by sb)"

Sometimes, determining whether the [g] sound of a [gwV] sequence is from **g** or **r** is difficult, especially when it is in the morpheme-internal position. For example, for the four cases of (84), a clue is reference to the Runyankore cognates. "Alcoholic drink" (84a) is [amáːgwa] in Rukiga and [amáːɾwa] in Runyankore. Because Runyankore does not pronounce **rwV** as [gwV], we can adopt the spelling **amáːrwa**. The Runyankore-Rukiga orthography also puts it as **amaarwa**. The same can be said of **endurwe** [endugwe] "bile" (84b). However, with **ekitákangurírwa** [ecitákaˈŋguɾígwa] "spider" (84c), the Runyankore word differs from the Rukiga word, and we cannot decide[19] even though the [igwa] sequence seems to be from the applicative suffix **-ir-** followed by the passive suffix **-w-**.

(84)	a.	amáːrwa 6	[amáːgwa]	"alcoholic drink"
		cf. Ank. *amáːrwa* 6	[amáːɾwa]	"alcoholic drink"
	b.	endurwe 9	[endugwe]	"bile"
		cf. Ank. *endurwe* 9	[endɾwe]	"bile"
	c.	ekitákangurírwa	[ecitákaˈŋguɾígwa]	"spider"
		cf. Ank. *orutáːngura* 11	[oɾutáːŋguɾa] or *ekitáːngura* 7	[ecitáːŋguɾa] "spider"

[19] In Rutooro, "spider" is **enyamumbúbi** 9, and in Runyoro, it is **enyamumbûbi** 9.

d.	okurwana	[okugwaˈna]	“to fight”
	cf. Ank. *okurwana*	[okuɾwaˈna]	“to fight”

Verbs whose radical begins with **r**, such as **okurwana** /o-ku-rwan-a/ [okugwaˈna] “to fight” (84d), cause a spelling problem when the nasal **n**-, the subject or object prefix of the first-person singular meaning “I” or “me,” is prefixed to them, as in (85).

(85) a. ngwáːna
n-rwan-a
1st.Per.Sg.SubPr-fight-FV
“I usually fight”

b. okungwanáho
o-ku-n-rwan-a hó
Aug15-NPr15-1st.Per.Sg.ObjPr-fight-FV Clit
“to help me in difficulty”

We know that the radical of the verb **okurwana** /o-ku-rwan-a/ [okugwaˈna] “to fight” begins with **r**. However, we hesitate to spell the sentence “I usually fight” as **nrwáːna** with **r** because **r** is preceded by **n**. Can we then put it as **ndwáːna** with **d**? We think it more appropriate to spell it **ngwáːna** using **g**.

Finally, an interesting situation is presented in (86). In Rukiga, “to be lucky” (86a) and “to be hunted (by sb)” (86b) become homonymous because **okuhîːrwa** “to be lucky” is pronounced as [okuhîːgwa], exactly the same as **okuhîːgwa** [okuhîːgwa] “to be hunted (by sb).” **Okuhîːrwa** [okuhîːgwa] ‘to be lucky’ looks like a passive form, but Rukiga does not have a related verb like **okuhîːra**.[20]

(86) a.	okuhîːrwa	[okuhîːgwa]	“to be lucky”
	cf. Ank. *okuhîːrwa*	[okuhîːrwa]	“to be lucky”
b.	okuhîːgwa	[okuhîːgwa]	“to be hunted (by sb)”
	cf. *okuhîːga*	[okuhîːga]	“to hunt”

3.17. mwV > mᵍwV > mᵍV > mŋV

The **mwV** sequence gives rise to **mŋV**. First, it becomes **mᵍwV** by velarization of **w**. Then, **w** drops, and the resulting **g** assimilates to the preceding **m** in terms of nasality, thus becoming **ŋ**. Because the nominal prefixes of class 1 and class 3 are **mu**- and the subject prefix of the second-person plural is **mu**-, **mŋV** is regularly seen in these forms

[20] The verb **okuhîːra** exists in Rukiga, but it means “to be cooked (ready) at a place.” It is an applicative form of **okúsa** (or **okúsya**) “to be cooked [of food].”

when **mu-** is followed by a vowel, as in (88). Additionally, **mŋV** forms are regularly seen in passive forms, as in (89).

(87) a. mwa > mŋa
b. mwe > mŋe
c. mwi > mŋi
d. mwo > mŋo
e. mwu > mŋu

(88) a.	omwâːna 1	[omŋâːna]	"child"
b.	omweyo 3	[omŋeˑjo]	"broom"
c.	mwine *2nd.per.pl.pres.*	[mŋiˑne]	"you have"
d.	omwóːjo 1	[omŋóːʒo]	"boy"
e.	omwûːya 3	[omŋûːja]	"air"
(89) a.	okúkoma	[okúkoma]	"to tie"
	okúkomwa *pass*.	[okúkomŋa]	"to be tied (by sb)"
b.	okútuma	[okútuma]	"to send (a person)"
	okútumwa *pass*.	[okútumŋa]	"to be sent (by sb)"
c.	okusiːma	[okusiːma]	"to thank"
	okusiːmwa *pass*.	[okusiːmŋa]	"to be thanked (by sb)"
	asíːmwe *3rd.per.sg.pass.subj*.	[asíːmŋe]	"(that) he/she be thanked (by sb)"

3.18. nwV > n^{g}wV > n^{g}V > nŋV

The sequence **nwV** becomes **nŋV**. The process is similar to that of **mwV** in Section 3.17. First, **nwV** becomes **n^{g}wV** by velarization of **w**. Then, **w** drops, and the resulting **g** assimilates to the preceding **n** in terms of nasality, thus becoming **ŋ**. Examples are not numerous, except for passive forms (92), in which case the **nwa** [nŋa] sequence regularly appears.

(90) a. nwa > nŋa
b. nwe > nŋe
c. nwi > nŋi (no example)
d. now > nŋo (no example)
e. nwu > nŋu (no example)

(91) a.	énwa	[énŋa]	"species of wasp"
b.	omujúnwa	[omuʒúnŋa]	"saved person"
(92) a.	okukona	[okukona]	"to castrate (small animals)"
	okukonwa *pass*.	[okukonŋa]	"to be castrated (by sb)"
b.	okúshana	[okúʃana]	"to repair (a house, etc.)"
	okúshanwa *pass*.	[okúʃanŋa]	"to be repaired (by sb)"

c.	okúbona	[okúβona]	"to find"
	okúbonwa *pass*.	[okúβonŋa]	"to be found (by sb)"
d.	zibónwe 10 *pass.subj.*	[ziβónŋe]	"(that) they be found (by sb)"

3.19. nywV (> nyᵍwV > nyᵍV > nyŋV)

With respect to the **nywV** sequence, pronunciation without the effect of velarization is possible, that is, **nywV** [ɲwV]. This means that the velarization effect by **w** is weak for **nywV** sequences. When velarization works, the process is similar to **mwV** sequences in Section 3.17 and **nwV** sequences in Section 3.18. That is, first, **nywV** becomes **nyᵍwV** by velarization of **w**. Then, **w** drops, and the resulting **g** assimilates to the preceding **ny** in terms of nasality, thus becoming **ŋ**. Examples are not numerous, except for passive forms, in which case the **nywa** [ɲwa] sequence regularly appears as in (95), and for subjunctive forms whose ending is the vowel **-e**, as in (94b).

(93)
a.	nywa	>	nyŋa	
b.	nywe	>	nyŋe	
c.	nywi	>	nyŋi	
d.	nywo	>	nyŋo	(no example)
e.	nywu	>	nyŋu	(no example)

(94)
a.	okúnywa	[okúɲwa]	or [okúɲŋa]		"to drink"
b.	ónywe *2nd.per.sg.subj.*	[óɲwe]	or [óɲŋe]		"(that) you drink"
c.	omúnywi 1	[omúɲwi]	or [omúɲŋi]		"drunkard"

(95)
a.	okutîːnya	[okutîːɲa]			"to fear"
	okutîːnywa *pass*.	[okutîːɲwa]	or	[okutîːɲŋa]	"to be feared (by sb)"
b.	okushêːnya	[okuʃêːɲa]			"to demolish"
	okushêːnywa *pass*.	[okuʃêːɲwa]	or	[okuʃêːɲŋa]	"to be demolished (by sb)"
c.	okumanya	[okumaŋa]			"to know"
	okumanywa *pass*.	[okumaŋwa]	or	[okumaŋŋa]	"to be known (by sb)"
d.	bimánywe 8 *pass.subj.*	[bimáŋwe]			"(that) they be know (by sb)"

3.20. ywV

The sequence **ywV** appears especially when verbs with the final radical **y** are followed by the passive suffix **-w-**, as in (96). However, no velarization effect to produce the consonant **g** is seen.

(96)
a.	okúgoya	[okúgoja]	"to knead (stiff porridge)"
	okúgoywa *pass.*	[okúgojwa]	"to be kneaded (by sb)"

	bugóywe 14 *pass.subj*.	[βugójwe]	"(that) it be kneaded (by sb)"
b.	okuramya	[okuɾaŋa]	"to worship"
	okuramywa *pass*.	[okuɾaŋwa]	"to be worshipped (by sb)"

4. Summary

This paper overviewed palatalization and velarization in Rukiga pronunciation. Because of the common Runyankore-Rukiga orthography, characteristics of Rukiga pronunciation are usually overlooked. Significantly, most writers of Runyankore-Rukiga are Runyankore speakers or Rukiga speakers influenced by Runyankore. Rarely is genuine Rukiga pronunciation brought to light. As emphasized elsewhere (see Kaji 2004 and 2023), elucidating minute features of languages proposes to preserve and testify to the heritage of the language and culture. Differences between languages embody their history and provide us precious data for reconstructing their history and, eventually, the history of the people who speak them.

The main points of palatalization in Rukiga pronunciation are recapitulated in (97a) and of velarization in (97b). According to the quality of **C** and **V**, various pronunciations are produced. They are quite characteristic when we hear them, but we do not realize this when we see them in dictionaries.

(97) a. **CyV** sequences generally produce **C^{t}yV** (where **C** is voiceless) or **C^{d}yV** (where **C** is voiced) sequences by palatalization of **y**, except for some cases in which spirantization works.

b. **CwV** sequences generally produce **C^{k}wV** (where **C** is voiceless) or **C^{g}wV** (where **C** is voiced) sequences by velarization of **w**.

Another important point about palatalization and velarization is that they involve a fortition process of consonants (98a,b). Seen in this light, understandably, **g** does not undergo either palatalization or velarization because it is a strong stop consonant [g]. The process in (98c) is excluded.

(98) a. β → b (see Sections 2.1 and 3.1)
b. r → d (see Section 2.16)
c. *γ → g

Acknowledgements

This is a product of research activity of the Institute for Comprehensive Research, Center for Language Studies, which was financially supported by the Kyoto Sangyo University Research Grant K2001. The travel expenses were funded by JSPS KAKENHI, Grant numbers 20K00556 and 24K03902.

References

Kaji, Shigeki. 2004. *A Runyankore Vocabulary*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Kaji, Shigeki. 2023. *A Rukiga Vocabulary.* Kyoto: Shoukadoh for the Center for Language Studies, Kyoto Sangyo University.

Lewis, M. Paul, Simons, Gary F. and Charles D. Fennig (eds.) 2015. *Ethnologue: Languages of the World*. 18th edition. Dallas, Texas: SIL International.

Makerere University. 2007. *Kashoboorozi y'Orunyankore-Rukiga*. Kampala: Fountain Publishers.

Mpairwe, Y., and Kahangi G.K. 2013. *A Learners' Runyankore-Rukiga Dictionary*. Kampala: Fountain Publishers.

Museveni, Yoweri K., Muranga, Manuel J.K., Gumoshabe, Gilbert, and Muhoozi, Alice N.K. 2012. *Katondoozi y'Orunyankore-Rukiga / Thesaurus of Runyankore-Rukiga*. Kampala: Fountain Publishers.

Tone patterns of Rukiga nominals: An overview

Shigeki Kaji
Kyoto Sangyo University

Abstract
This paper overviews the tonal patterns of Rukiga nominals, providing all the patterns yet found by us (no previous studies), and attempts to analyze them from a typological view. Rukiga is a Bantu language spoken in the southwestern region of Uganda. It is closely related and similar to its northern neighbor Runyankore, and the two languages are usually treated together as Runyankore-Rukiga. However, there is a significant difference between Rukiga and Runyankore in tonal patterns of nominals because Rukiga has nominals with two high tones in a word in isolation, but Runyankore does not. Moreover, concerning tone rules, High Tone Anticipation works similarly in Rukiga and Runyankore, but in Rukiga, it works more strongly than Runyankore such that the high tone of not only the final but also the penultimate syllables of words is anticipated in isolation. Tone Reduction rules that work in Runyankore do not work in Rukiga. This paper reveals the tonal features of Rukiga, which have been overshadowed by Runyankore.

Keywords: Rukiga, Runyankore, Bantu, tone pattern, High Tone Anticipation

1. Introduction

Rukiga is a Bantu language spoken by 1,580,000 speakers (Lewis *et al*. 2015) in the southwestern corner of Uganda. This language is similar to a language to the north, Runyankore. The two languages are commonly treated together under the name Runyankore-Rukiga. Rukiga has a common orthography with Runyankore. However, the pronunciation of Rukiga sometimes substantially differs from that of Runyankore, and the differences are veiled by the common orthography.[1]

This paper overviews the tonal patterns of Rukiga nominals, referring to Runyankore when necessary or relevant.[2] Rukiga can roughly be divided into two dialects: the northern dialect and the southern dialect. The main data in this paper are from the southern dialect, spoken around Lake Bunyonyi, considered more genuine, retaining archaic features, than the northern dialect. The northern dialect is spoken south of Runyankore

[1] Segmental differences are presented in Kaji (2023, 2024).
[2] All the data used in this paper are from the author's fieldwork. Part of his research results on Runyankore is published as Kaji (2004, 2010). His Runyankore data are from the dialect of Ntungamo District. Rukiga data in the form of vocabulary is published as Kaji (2023).

and has features in common with Runyankore. In Section 6, tonal characteristics of the northern dialect are discussed.

Nominals are forms that take a nominal prefix and are mostly nouns. The examples used in this study are all nouns, but some qualificative adjectives (like "good" or "bad") that take a nominal prefix are also in the category. Other forms, namely, pronominal forms (i.e., demonstratives, connectives, relative forms, etc.), numeral forms, and verbal (non-relative) forms, are not dealt with in this study. They share tonal features with nominals, but other factors intervene in them.

2. Tone value

Rukiga has two basic tones, high (H) and low (L).[3] H is indicated with an acute accent mark over the vowel (**á**), and L is unmarked. Falling tone (F), marked with a circumflex over the vowel (**â**), appears. Rising tone (R) does not appear. In this section, two phonetic features of Rukiga tones are explained, one regarding L tone, and the other regarding H tone.

2.1. Lowering of low tone

Words and phrases are pronounced on a slightly descending pitch toward the end in Rukiga. Particularly, utterance-final L tone is realized lower than normal L. This phenomenon is especially notable in sequences of consecutive L-toned syllables, as in (1). For example, **ekicu** "clouds" is pronounced low-low-lower in utterance-final position. This lowing of L tone at the end is remediated when pause is eliminated by modifying the word with another word (e.g., adjective) that directly follows the noun. See (2).

(1)	a.	ekicu[4]	[_ _ _]	"clouds"
	b.	okuguru	[_ _ _ _]	"leg"
	c.	orushahuzo	[_ _ _ _ _]	"curved machete"
	d.	okujuba	[_ _ _ _]	"to get wet, getting wet"
(2)	a.	ekicu kyangye	[_ _ _ _ _]	"my clouds"
	b.	okuguru kwangye	[_ _ _ _ _ _]	"my leg"
	c.	orushahuzo rwangye	[_ _ _ _ _ _ _]	"my curved machete"
	d.	okujuba kwangye	[_ _ _ _ _ _]	"my getting wet"

[3] We do not consider the theoretical question of whether the opposition is either H vs. L or H vs. Ø. Simply, H and L are used. In this paper, the term "tone" is used both phonologically (as opposed to accent) and phonetically.

[4] The orthographical spelling is used in this paper to describe Rukiga and Runyankore words. The following conventions are applied: **bb** [b], **b** [β] (but **mb** [mb]), **ki** [ci], **kye** [ce], **gi** [ɟi], **gye** [ɟe], **sh** [ʃ], **j** [ʒ], **ny** [ɲ], and **y** [j]. The letter **c** represents [c] in Rukiga but [ʧ] in Runyankore. The length mark **:** and tone markings ´ and ˆ were added by us to the orthographical spelling.

2.2. H tone as F tone

H tone does not appear every time as H. Characteristically, it becomes F in penultimate heavy syllables in utterance-final position, as in (3).[5] This F tone is rectified to H when the noun is modified with an adjective that follows it, eliminating the pause after the noun, as in (4).

(3)	a.	ekibâːju	[_ _ ↘ _]	"blister"
	b.	enyonyôːzi	[_ _ ↘ _]	"star"
	c.	ekyâːra	[_ ↘ _]	"nail"
	d.	akakyâːmba	[_ _ ↘ _]	"grassland"
(4)	a.	ekibáːju kyangye	[_ _ – _ _ _]	"my blister"
	b.	enyonyóːzi yangye	[_ _ – _ _ _]	"my star"
	c.	ekyáːra kyangye	[_ – _ _ _]	"my nail"
	d.	akakyáːmba kangye	[_ _ – _ _ _]	"my grassland"

Notably, F tone appears only on the heavy penultimate syllable of words. It never appears on the antepenultimate or more left syllable of words, even though the syllable has a long vowel. It does not appear on the final syllable of words either because the final syllable of words does not accommodate two moras in Rukiga. If a resulting F tone is not in the penultimate syllable of words, it is realized as H.

(5) indicates how Rukiga treats F tone that occurs because of the concatenation of H and L. In (5a), **ní ente** H LL first becomes **n'êːnte** FL, but because F tone is not permitted as a basic tone, it is recognized as H. This H tone is on a long penultimate syllable; thus, it is realized as F in the last place. In (5b), the resulting long vowel [îː] from **ní ibega** H LLL is recognized as H, as in (5a). Because it is in the antepenultimate syllable, it cannot be F but is realized as H. In (5c) **ní omúshure**, H LHLL first becomes **n'ômúshure** by vowel elision. Notably, the resulting vowel is short. This resulting F tone is recognized as H because it is in the preantepenultimate syllable where F cannot be accommodated.

(5)	a.	egi	ní	ente	→ egi n'êːnte	→ egi n'éːnte → egi n'êːnte
		this	is	cow		
		"This is a cow."				
	b.	eri	ní	ibega	→ eri nîːbega	→ eri níːbega
		this	is	shoulder		

[5] Heavy syllables mean syllables with a long vowel. Long vowels are either phonemic or phonetic. The long vowels [âː] in **ekibâːju** "blister" (3a) and [ôː] in **enyonyôːzi** "star" (3b) are phonemic. The long vowels [âː] in **ekyâːra** "nail" (3c) and in **akakyâːmba** "grassland" (3d) are phonetic. The phonetic long vowel [âː] in **ekyâːra** "nail" emerges by glide formation and compensatory lengthening, and that of [âː] in **akakyâːmba** "grassland" is caused by the following nasal cluster. Long vowels, especially phonetic long vowels, do not become two mora-long unless their tone is H or F, in which case the length mark **ː** is not used in this paper.

"This is a shoulder."

c. ogu ní omúshure → ogu n'ômúshure → ogu n'ómúshure
this is mouse bird
"This is a mouse bird."

3. Noun classes

This section briefly checks the morphological structure of nominals because we cannot discuss the tonal patterns of nominals without knowing their morphological structure. Rukiga nouns are divided into 21 classes, numbered from 1 to 21. In what follows, the class number is indicated after the noun in question. Usually, two classes, such as 1–2 and 3–4, are paired to denote the singular and the plural of a noun; however, other patterns exist, such as classes 9, and 10, which have the same form for both the singular (class 9) and the plural (class 10).

Morphologically, the class of a noun is indicated by its nominal prefix. (6) is the morphological structure of nominal forms. There are four types. (6a) is a complete form comprising three elements: the augment, nominal prefix, and stem. The augment (or initial vowel) has an article-like function, the use of which depends on the syntactic environment.[6] (6b) is a form that has only an augment and a stem, with a zero nominal prefix. (6c) is a form that has a nominal prefix and a stem. This form has a zero augment. In (6d), only the stem is observed. This form has a zero augment and a zero nominal prefix. Type (6a) is fundamental and typical and has the most examples among the four types.

(6) a. augment-nominal prefix-stem
ex. o-mu-zi 3, e-mi-zi 4 "root"

b. augment-Ø-stem
ex. e-Ø-shô꞉nda 9,10 "corner"

c. Ø-nominal prefix-stem
ex. Ø-í-tama[7] 5 "cheek"
ex. Ø-za꞉-burangíti 10*b*[8] "blankets" < Sw.[9] *blanketi*

d. Ø-Ø-stem
ex. Ø-Ø-burangíti 9,10 "blanket" < Sw. *blanketi*

[6] For example, the augment is not used when the noun is the object of a negative transitive verb.
[7] The plural is **a-má-tama** 6 and has the three elements.
[8] 10*b* means this noun is a subtype of class 10 with its **za꞉-** prefix. In the examples below, 10*a* words also appear. 10*a* words are another subtype of class 10 nouns with a zero prefix.
[9] The symbol < means "borrowing." The abbreviations for language names are the following: Sw. for Swahili, Eng. for English, and Rw. for Kinyarwanda (Rwandan).

4. Tone patterns

From (7) to (11), we list all tone patterns found in nominals according to the length of the stem. A hyphen is placed between the nominal prefix and the stem. Each pattern is illustrated with two nouns if we find more than two examples. The stem is from one to five syllables long. No nominals of six or more syllables long are found except for some compounds and reduplicated forms. The pronunciation of the words of the lists is that in isolation, that is, when words are pronounced utterance-finally, which does not necessarily represent the underlying form. Therefore, the underlying H-toned syllable is underlined. Tonal patterns indicated in capital letters are underlying patterns of the stems. We explain underlying and surface tones in Section 5.1.

(7) one-syllable stem nouns

L	omu-zi 3, emi-zi 4	"root"
	omu-mwa 3, emi-mwa 4	"mouth"
H	omú-ti 3, emí-ti 4	"tree"
	okú-jwi 15, amá-jwi 6	"knee"

(8) two-syllable stem nouns

LL	oku-guru 15, ama-guru 6	"leg"
	oru-tindo 11, en-tindo 10	"bridge"
LH	eki-sháka 7, ebi-sháka 8	"bush"
	eki-júma 7, ebi-júma 8	"fruit"
HL	omú-tima 3, emí-tima 4	"heart"
	orú-kumu 11, én-kumu 10	"finger"

(9) three-syllable stem nouns

LLL	oru-shahuzo 11, en-shahuzo 10	"curved machete"
	omu-shoko:ro 3, emi-shoko:ro 4	"bone marrow"
HLH	aka-gyégyéna 12, obu-gyégyéna 14	"strainer"
	en-tó:ndóre 9,10	"fresh ripe beans"
LHL	eki-tágata 7, ebi-tágata 8	"hot spring"
	eny-ó:ngyeza[10] 9,10	"added bonus"
HLL	omu-nyúruru 3, emi-nyúruru 4	"door handle"
	oru-zítiro 11, en-zítiro 10	"fence, hedge"

(10) four-syllable stem nouns

LLLL	eki-huruguto 7, ebi-huruguto 8	"earwax"
	omu-gambirizi 3, emi-gambirizi 4	"speaker"
LLLH	e-kerezíya 9,10	"Catholic church" <Eng. *ecclesia*

[10] This word may be analyzed as **e-nyó:ngyeza** 9,10, that is, **ny-** as part of the stem. Neither analysis affects our discussion. The same can be said about **ny-irá:nsáse** 9,10 "variety of sweet potato" in (10).

	burangíti 9*a*,10*a*, za:-burangíti 10*b*	"blanket" <Sw. *blanketi*
LHLH	ny-irá:nsáse 9,10	"variety of sweet potato" <Rw.
HLHL	omu-kárabâ:nda 3, emi-kárabâ:nda 4	"wooden case"
	en-górofâ:ni 9,10	"wheelbarrow"
HLLH	omu-kóroró:mbya 3, emi-kóroró:mbya 4	"rainbow"
	omu-bírikíra 3, emi-bírikíra 4	"funnel"
LLHL	en-tangahû:zi 9,10	"ginger" <Sw. *tangawizi*
LHLL	eki-korógoto 7, ebi-korógoto 8	"dried papyrus plant"
	aka-samúnyiga 12, obu-samúnyiga 14	"skunk (?)"
HLLL	omu-cúraguzi 1, aba-cúraguzi 2	"night dancer"
	aka-jágararo 12, obu-jágararo 14	"public excitement"

(11) five-syllable stem nouns

LLLLL	musingiriro 9,10	"variety of bean"
LLLLH	kabaragára 9,10	"small sweet banana"
	karimagézi 9*a*,10*a*, za:-karimagézi 10*b*[11]	"computer"
LLHLL	pa:sikárama 9,10	"lawn"
HLHLL	en-dábarásanyi 9,10	"species of weed, hairy beggar-ticks"
HLLLL	omw-é:yayamuro 3, emy-é:yayamuro 4	"yawn"

In addition to those listed above, some words like those in (12) are observed, which have two tone patterns.

(12)	a.	HL	amá-gezi 6	"practical wisdom, quick-wittedness"
		or LH	ama-gézi 6	
	b.	LL	oru-taro 11, en-taro 10	"war, battle"
		or LH	oru-táro 11, en-táro 10	
	c.	HLL	eki-bíri:ti 7, ebi-bíri:ti 8	"matchbox" <Sw. *kiberiti*
		or LHL	eki-birî:ti 7, ebi-birî:ti 8	

Finally, some minimal pairs by tone are presented in (13).

(13)	a.	L	omu-ra 3, emi-ra 4	"rectum"
		H	omú-ra 3, emí-ra 4	"flea"
	b.	LL	em-piri 9,10	"wooden mallet"
		HL	ém-piri 9,10	"puff adder"
	c.	LH	eki-sísi 7, ebi-sísi 8	"stump"
		HL	ekí-sisi 7, ebí-sisi 8	"gourd"

[11] This is a compound word containing **amagézi** 6 "practical wisdom."

d. LL en-kombe[12] 9,10 "protruding forehead"
LH en-kó꞉mbe 9,10 "millet porridge"
HL en-kô꞉mbe 9,10 "pigeon, dove"

5. Analysis

In examples (7)-(11), three observations, as in (14), are made about tone, and discussed in Sections 5.1, 5.3, and 5.4, respectively.

(14) a. Generally, the nominal prefix, and the augment are not H-toned. The exceptions are words of H and HL stems.
b. Most words can have only one H tone in them, but some words have two H tones. No word is found to have three or more H tones.
c. Although Rukiga has various tone patterns that neighboring languages do not have, its system is not unrestricted (which allows any combinations of H and L in a word).

5.1. Underlying and surface tones

In Rukiga, pause influences pronunciation, and the underlying tone of words does not appear as such in isolation. Two phenomena have been mentioned: lowering of L tone (Section 2.1) and H tone realized as F (Section 2.2). This section deals with High Tone Anticipation, which has a significant impact on the realization of H tones.

5.1.1. High tone anticipation

High Tone Anticipation (HTA) is a widespread rule in Bantu languages of the area, including Rukiga, and Runyankore. Notably, it applies more strongly in Rukiga than in the aforementioned languages. In (14a), we said: "Generally, the nominal prefix, and the augment are not H-toned. The exceptions are words of H and HL stems." HTA explains these exceptions.

The examples in (15) are retaken from (7) and (8), in which the possessive adjective **-angye** "my" is added to modify the noun, eliminating pause. The possessive adjective is after the noun it modifies. We thereby confirm the original position of H tone for each word.

For example, the word for "tree" is pronounced **omúti** in isolation, but **omutí** when it is followed by **gwangye**[13] "my" (15a). That is, the H tone is originally on the last

[12] The vowel **o** of **enkombe** is long but not fully two-mora long ([eŋkoˑmbe]), in which case the length mark **꞉** is not used. See Note 5.
[13] The possessive adjective stem **-angye** "my" takes a prefix **gu-** (cl.3), **ki-** (cl.7), **ru-** (cl.11), **ku-** (cl.15), etc. according to the class of the noun it modifies. The prefixes that possessive adjectives take are pronominal but similar in form to nominal prefixes.

syllable of the word and is anticipated by one syllable in isolation. To make explicit the location of underlying H, we underline underlying H-toned syllables. For words of L and LL stems, which have no H tone, no such tone shift occurs. Only words of H, LH, and HL stems undergo HTA. We verify in (15) that each H tone is anticipated by one syllable in isolation. This is why words of H and HL stems have H tone on the noun prefix in isolation, which is shifted from right to left by one syllable. Thus, the noun prefix (and the augment) is L-toned underlyingly.

(15) a. one-syllable stem nouns

L	omu-zi 3	"root"
	omu-zi gwangye 3	"my root"
H	omú-ti 3	"tree"
	omu-tí gwangye 3	"my tree"

b. two-syllable stem nouns

LL	oku-guru 15	"leg"
	oku-guru kwangye 15	"my leg"
LH	eki-sháka 7	"bush"
	eki-shaká kyangye 7	"my bush"
HL	omú-tima 3	"heart"
	omu-tíma gwangye 3	"my heart"

Notably, in the examples in (15), the H tone of the last and penultimate syllables of words is anticipated. In **omú-ti** "tree" and **eki-sháka** "bush," the H tone of the last syllable is anticipated, and in **omú-tima** "heart," the H tone of the penultimate syllable is anticipated. The HTA of the penultimate syllable of words is characteristic of Rukiga tone. In other languages in the area, including Runyankore, HTA occurs only of H tone of the last syllable of words, as we present in Section 5.4.[14]

Also important is that the H tone in the antepenultimate or more left syllable of a word does not move no matter what word is before or after the word. This can be verified in (16). For example, in **aka-gyégyéna** "strainer" and **omu-nyúruru** "door handle," the H tone of the antepenultimate syllable remains in the original position in isolation. In **omu-kárabâ:nda** "wooden case" and **omu-cúraguzi** "night dancer," the H tone of the preantepenultimate syllable remains in the original position in isolation.

[14] In the northern dialect of Rukiga, which shares many features with Runyankore, HTA occurs only of H on the final syllable of words as Runyankore. This point is exemplified in Section 6.

(16) a. three-syllable stem nouns

LLL	oru-shahuzo 11	“curved machete”
	oru-shahuzo rwangye 11	“my curved machete”
HLH	aka-gyégyéna 12	“strainer”
	aka-gyégyená kangye 12	“my strainer”
LHL	eki-tágata 7	“hot spring”
	eki-tagáta kyangye 7	“my hot spring”
HLL	omu-nyúruru 3	“door handle”
	omu-nyúruru gwangye 3	“my door handle”

b. four-syllable stem nouns

LLLL	eki-huruguto 7	“earwax”
	eki-huruguto kyangye 7	“my earwax”
LLLH	e-kerezíya 9	“Catholic church”
	e-kereziyá yangye 9	“my Catholic church”
LHLH	ny-irá:nsáse 9	“variety of sweet potato”
	ny-irá:nsasé yangye 9	“my variety of sweet potato”
HLHL	omu-kárabâ:nda 3	“wooden case”
	omu-kárabá:nda gwangye 3	“my wooden case”
HLLH	omu-kóroró:mbya 3	“rainbow”
	omu-kórorombyá gwangye 3	“rainbow”
LLHL	en-tangahû:zi 9	“ginger”
	en-tangahú:zi yangye 9	“my ginger”
LHLL	eki-korógoto 7	“dried papyrus plant”
	eki-korógoto kyangye 7	“my dried papyrus plant”
HLLL	omu-cúraguzi 1	“night dancer”
	omu-cúraguzi wangye 1	“my night dancer”

c. five-syllable stem nouns

LLLLL	musingiriro 9	“variety of bean”
	musingiriro yangye 9	“my variety of bean”
LLLLH	kabaragára 9	“small sweet banana”
	kabaragará yangye 9	“my small sweet banana”
LLHLL	pa:sikárama 9	“lawn”
	pa:sikárama yangye 9	“my lawn”
HLHLL	en-dábarásanyi 9	“species of weed, hairy beggar-ticks”
	en-dábarásanyi yangye 9	“my hairy beggar-ticks”
HLLLL	omw-é:yayamuro 3	“yawn”
	omw-é:yayamuro gwangye 3	“my yawn”

The above demonstrates that the H tone of the last and penultimate syllables of a

word is anticipated by one syllable in isolation. However, there is one condition for this rule to apply: the original H-toned syllable must be a light one. That is, its vowel must be short. If the vowel is long, HTA does not occur. In **omu-kárabâ:nda** 3 "wooden case" and **en-tangahû:zi** 9 "ginger" of (16b), the H tone of the penultimate syllable does not move.

In this respect, the word for "bullet" is notable. It has two pronunciations, either **isâ:si** 5 (17a) or **ísasi** 5 (17b) in isolation. Both pronunciations have the same underlying stem pattern, HL. The difference is whether the vowel of the penultimate syllable of the word is long (**isâ:si** 5) or short (**ísasi** 5). In **isâ:si** 5 (17a), because the vowel of the penultimate syllable is long, HTA does not apply. In **ísasi** 5 (17b), by contrast, because the vowel of the penultimate syllable is short, HTA applies. When the word is modified by the adjective **-angye** "my," the tone pattern of the stem emerges as HL for both pronunciations.

(17) a. isâ:si 5, amasâ:si 6 "bullet" < Sw. *risasi*
isá:si ryangye 6 "my bullet"
b. ísasi 5, amásasi 6 "bullet" < Sw. *risasi*
isási ryangye 5 "my bullet"

Finally, HTA occurs in other categories of words than nouns. In (18a), for example, the H tone of the clitic **kí** "what (kind of)" is transferred to the last syllable of the word **omuga:ti** "bread." In (18b), the H tone of **túso** "like us," on the penultimate syllable of the word, is transferred to the last syllable of the preceding word. The left word **tukoraga** "we always work" is a conjugated verb form.[15]

(18) a. Ogu ní omuga:ti kí? → Ogu n'ómuga:tí ki? "What (kind of) bread is this?"
b. Tukoraga túso. → Tukoragá tuso. "We always do like that."

5.1.2. Transferred opposition

We confirmed in the previous section that the H tone of both the last and penultimate syllables of words is anticipated by one syllable in isolation unless the vowel of the original H-toned syllable is long. As a corollary of this Rukiga HTA, we understand that the underlying opposition L-LH vs. L-HL is transferred to the opposition L-HL vs. H-LL in isolation, as in (19).

(19) a.	L-LH	→	L-HL	utterance-finally
	ex. eki-shaká		eki-sháka	"bush"

[15] When the present habitual verb form "we always work" is pronounced in isolation, it is **tukorága**. However, when it is followed by another word in a phrase, its H tone disappears, as in **Tukoraga omurimo.** "We always do a/the work." This is a case of conjoint/disjoint alternation.

b.	L-HL	→	H-LL	utterance-finally
	ex. omu-tíma		omú-tima	"heart"

For the opposition of the underlying -LHL vs. -HLL, the opposition is neutralized in isolation, both being -HLL, because the antepenultimate H tone is not anticipated in isolation, as in (20b).

(20)	a.	-LHL	→	-HLL	utterance-finally
		ex. eki-tagáta		eki-tágata	"hot spring"
	b.	-HLL	→	-HLL	utterance-finally
		ex. omu-nyúruru		omu-nyúruru	"door handle"

In the case of the four-syllable stem words, the underlying opposition -LLHL vs. -LHLL (even though the opposition is that of penultimate vs. antepenultimate as in (20)), is maintained as such in isolation, as in (21). This is because the vowel of the penultimate syllable of **en-tangahú:zi**9,10 "ginger" is long, and its H tone is not anticipated. Regarding whether syllable length, namely, three-syllable long vs. four-syllable long, must be taken into consideration for HTA to apply, we need to withhold a conclusion because **en-tangahú:zi** 9,10 "ginger" is the sole example of -LLHL stem and is a borrowing from Swahili.

(21)	a.	-LLHL	→	-LLHL	utterance-finally
		ex. en-tangahú:zi		en-tangahû:zi	"ginger" < Sw. *tangawizi*
	b.	-LHLL	→	-LHLL	utterance-finally
		ex. eki-korógoto		eki-korógoto	"dried papyrus plant"

In summary, we present (22) regarding the tonal pronunciation of nominals in isolation.

(22) a. If the penultimate syllable of a word is H-toned (its vowel is either short or long), it is the final syllable that is H-toned underlyingly.
For a short vowel: ex. eki-sháka "bush" (19a)
For a long vowel: ex. omu-kóroró:mbya "rainbow" (16b)

b. If the penultimate syllable of a word is F-toned (this automatically means its vowel is long), the penultimate syllable is H-toned underlyingly.
ex. en-tangahû:zi "ginger" (16b)

c. If the antepenultimate syllable of a word is H-toned (its vowel is either short or long), this indicates two possibilities, either the antepenultimate syllable is H-toned or the penultimate syllable is H-toned underlyingly and its H tone is

anticipated by one syllable.
For the former: ex. omu-nyúruru "door handle" (20b)
For the latter: ex. omú-tima "heart" (19b) and eki-tágata "hot spring" (20a)

5.2. High tone bumping by Hyman (2022)

Hyman (2022) proposed a rule called High Tone Bumping (HTB) in his analysis of Runyankore tone. Runyankore HTB is exemplified in (23a).[16] Because of the H tone of the interrogative clitic **kí** "what (kind of)," the H tone of **wá** of **ka:wá** "coffee" is "bumped" onto the preceding syllable. Subsequently, the H tone of **kí** is retracted to the vacant **wa**. This latter process is triggered by the inserted utterance-final L tone and Final High Retraction (FHR). FHR is almost the same as the present author's HTA except that in Rukiga, not only the final but also the penultimate H is anticipated onto the preceding syllable. The same method of analysis is applicable to the Rukiga cognate, as in (23b).

(23)	a.	ka:wá[17] kí?	→	ká:wá ki?	
		coffee what (kind of)		"what (kind of) coffee?"	
	b.	ka:wá kí?	→	ká:wá ki?	
		coffee what (kind of)		"what (kind of) coffee?"	

In this manner, HTB proves applicable in Rukiga, too. However, for HTB to work in Rukiga with a bumping word that begins with a vowel, one operation, namely, syllable reduction, must be applied beforehand. See (24).

(24)	a.	enjá:ngu 9,10			"cat"
	b.	enjangú[18] yangye 9			"my cat"
		cat my			
	c.	enjá:ngú émwe 9	→	enjá:ng'é:mwe	"one cat"
		cat one			
(25)	a.	underlying:		enjangú emwé	"one cat"
	b.	syllable reduction:		enjang'é:mwé	
	c.	HTB:		enjá:ng'é:mwe	

A cat is called **enjá:ngu** 9 in Rukiga. Its pronunciation in isolation is LHL (24a) because it has a final-H pattern underlyingly. When it is followed by the adjective **yangye** "my" in a noun phrase, the underlying pattern LLH emerges (24b). A problem occurs

[16] (23a) is the present author's representation with his data.
[17] **ká:wa** is a borrowing from Swahili.
[18] Although the length mark **:** is not used, **a** of **enjangú** is long (but not two mora long). See footnote 5.

when, for example, the adjective **émwe** "one" follows it, which begins with a vowel. In (24c), the sequence **enjá:ngú émwe** looks as if it has three H-toned syllables. In reality, the pronunciation is **enjá:ng'é:mwe**. That is, the final vowel of **enjá:ngú** is elided, and the number of syllables reduces by one. See (25b). After this operation, HTB applies to **enjang'é:mwé** to produce **enjá:ng'é:mwe**. The H tone of the elided **ú** remains on the eliding **e**.

HTB paired with syllable reduction works in other syntactic contexts. See (26).

(26)	a.	ítama 5, amátama 6			"cheek"
	b.	itáma ryangye 5 cheek my			"my cheek"
	c.	itáma rímwe 5 cheek one			"one cheek"
	d.	erí ní ítama this is cheek	→	erí ní:tama	"This is a cheek."
(27)	a.	underlying:		eri ní itáma	"This is a cheek."
	b.	syllable reduction:		eri ní:táma	
	c.	HTB:		erí ní:tama	

A cheek is **ítama** 5 in isolation. When it is followed by a consonant-initial word in a phrase, as in (26b) or (26c), the underlying pattern LHL emerges. However, in a sentence like "This is a cheek." (26d), cl.5 demonstrative **eri** LL "this" has an H on its final syllable. To understand this, we have to know that syllable reduction first applies to the sequence **ní itáma** to change it to **eri ní:táma** before HTB applies to it to make it **erí ní:tama**. See the process in (27).

5.3. Two H tones in one word

A noteworthy characteristic of Rukiga nominal tone is that some words have two H tones in a word in isolation. This is exemplified by words of HLH, LHLH, HLHL, and HLLH stems, retaken in (28) from (9) and (10).

(28)	a.	three-syllable stem nouns		
		HLH	aka-gyégyéna 12, obu-gyégyéna 14	"strainer"
			en-tó:ndóre 9,10	"fresh ripe beans"
	b.	four-syllable stem nouns		
		LHLH	ny-irá:nsáse 9,10	"variety of sweet potato" < Rw.
		HLHL	omu-kárabâ:nda 3, emi-kárabâ:nda 4	"wooden case"
			en-górofâ:ni 9,10	"wheelbarrow"

HLLH omu-kóroró:mbya 3, emi-kóroró:mbya 4 "rainbow"
omu-bírikíra 3, emi-bírikíra 4 "funnel"

We do not have a definite solution to the analysis of this category of words. However, no word of LLH pattern among three-syllable stem nouns is observed, though we find many final-H pattern words for two-syllable stem words. If this observation is correct, the HLH pattern is a realization of LLH. For some reason, the stem-initial (or antepenultimate) syllable of the word is realized as H-toned when the final syllable is also H-toned. In the same manner, the LHLH pattern of **ny-irá:nsáse** 9,10 "variety of sweet potato" may be considered a variation in LLLH. This word is special because it is a crop name[19] borrowed from Kinyarwanda. It is the sole word of the LHLH pattern of four-syllable stem nouns. We may put it on hold for the time being. Therefore, the remaining HLHL, and HLLH stem patterns present a more challenging situation than this.

Runyankore, the northern neighbor of Rukiga, does not have nominals of two H tones in one word in isolation (Kaji 2010). In this respect, comparing Rukiga words with two H tones to their Runyankore cognates is useful.

(29)	a.	Kiga	omunyórongóto 3, eminyórongóto 4	"worm"
		Nkore	omunyongorórwa 3, eminyongorórwa 4	
	b.	Kiga	akagyégyéna 12, obugyégyéna 14	"strainer"
		Nkore	akagyegyéna 12, obugyegyéna 14	
	c.	Kiga	aká:tíka 12, obwá:tíka 14	"broken piece"
		Nkore	aka:tíka 12, obwa:tíka 14	
(30)	a.	Kiga	egíríta 9,10	"razor"
		Nkore	egírita 9,10	
	b.	Kiga	ekishúshú:ngwa 7, ebishúshú:ngwa 8	"chaff, husks"
		Nkore	ekishúshungwa 7, ebishúshungwa 8	
(31)	a.	Kiga	ekánísa 9,10	"Protestant church" <Sw. *kanisa*
		Nkore	ekanísa 9,10	
	b.	Kiga	omuzígíti 3, emizígíti 4	"mosque" <Sw. *msikiti*
		Nkore	omuzigiti 3, emizigiti 4	

A number of words like those in (29) in which the left H tone of Rukiga words is not observed in the Runyankore cognates are found. This is noticeable. However, other patterns are observed. In (30), the situation is reversed. The Rukiga right H tone does not appear in Runyankore. In (31a), the location of neither of the Rukiga two H tones does

[19] Crop names are cultural words. They are coined in various ways and easy to diffuse.

not correspond to the Runyankore H tone. In (31b), the Rukiga word has two H tones, and the Runyankore cognate has no H tone.

Because things are not straightforward, we need to examine them one by one. There is one phenomenon that must be related to this point: the formation of negative verb infinitives. (32) presents verbs of L-toned radicals, and (33) presents verbs of H-toned radicals in Rukiga. The radical is bounded by a hyphen in the affirmative, and in the negative, the negative marker **-ta-** is bounded by a hyphen.

(32) L-toned radicals

a.	oku-shob-o꞉rokyerwa	→	okú-ta-shobó꞉rokyerwa
	"to understand"		"not to understand"
b.	oku-gar-ukani꞉sa	→	okú-ta-garúkani꞉sa
	"to reconcile"		"not to reconcile"
c.	oku-gyend-a	→	okú-ta-gyé꞉nda
	"to go"		"not to go"

(33) H-toned radicals

a.	oku-téb-enkani꞉sa	→	okú-ta-tebenkaní꞉sa
	"to prepare"		"not to prepare"
b.	oku-shém-ezibwa	→	okú-ta-shemezíbwa
	"to be operated"		"not to be operated"
c.	okú-boh-a	→	okú-ta-bóha
	"to bind"		"not to bind"

Notably, the addition of one element, the negative marker **-ta-**, gives rise to two H-toned syllables in one word. With L-toned radical verbs in (32), both the infinitive marker (class 15) **-ku-** and the right syllable of the radical are raised H. With H-toned radical verbs in (33), both the infinitive marker (class 15) **-ku-** and the rightmost syllable of the word are raised H, and the H tone of the radical disappears.[20] In (32c) and (33a,b,c), HTA applies in the negative. The examples in (32) show that the two H tones are actually from one, split into two. Because **-ku-** of both L-toned and H-toned radical verbs is raised H, its H tone may be additional,[21] not functioning to distinguish between L-toned and H-toned radical verbs. Runyankore does not raise it H in making the negative form of verb infinitives.[22] Although this is not a definite treatment of two H tones in one word, we do not take relevant the left H tone of two H tones for the time being.

One word has two pronunciations, with either two H tones or one H tone in a word.

[20] It is conceivable that the H tone of the radical moves to the rightmost syllable of the word.
[21] Conceivably, **-ta-** is H-toned and its H is anticipated onto the preceding **-ku-**.
[22] L-toned radical: **okugyenda** "to go" → **okutagyé꞉nda** "not to go"
H-toned radical: **okubóha** "to bind" → **okutáboha** "not to bind"

See (34a) and (34b).

(34) a. enkurúbure 9,10 "bare thing whose surface is removed"
enkurúbure yangye 9 "my bare thing whose surface is removed"
b. enkúrubúre 9,10 "bare thing whose surface is removed"
enkúruburé yangye 9 "my bare thing whose surface is removed"

The noun in (34) is derived from the verb **okukurubura** (with the L-toned radical -**kur-**) "to remove the surface to make it bare." Notably, in (33a), the deverbative process makes one H tone appear in the noun; by contrast, in (34b), two H tones appear in the noun: one on the stem-initial syllable **kú** and the other on the final syllable **ré** underlyingly. This also seems to suggest that if the rightmost syllable of the word is H, another syllable (the stem-initial, or the second, or preantepenultimate syllable of the word?) may also become H.

As aforementioned, regarding the four-syllable stem words of HLHL and HLLH patterns, we do not have a definite analysis. (35) is a retake of words of LLLH, HLHL, and HLLH stems with additional examples.

(35) a. words of LLLH stem
e-kerezíya 9,10 "Catholic church" < Eng. *ecclesia*
burangíti 9*a*,10*a*, za:-burangíti 10*b* "blanket" <Sw. *blanketi*
bbiringánya 9*a*,10*a*, za:-bbiringánya 10*b* "eggplant" < Sw. *biringanya*
b. words of HLHL stem
omu-kárabâ:nda 3, emi-kárabâ:nda 4 "wooden case"
en-górofâ:ni 9,10 "wheelbarrow"
c. words of HLLH stem
omu-kóroró:mbya 3, emi-kóroró:mbya 4 "rainbow"
omu-bírikíra 3, emi-bírikíra 4 "funnel"
obu-cúrigányi 14 "titillater"
aka-zé:ngyeréra 12, obu-zé:ngyeréra 14 "giddiness, dizziness"
i-húrungúru 5, ama-húrungúru 6 "dung of sheep or goats"
ama-shémerérwa 6 "happiness, joy"

Based on (35), the following observations are made. First, words of LLLH stems (35a) are few: we find only three examples, all borrowings. Second, words of HLHL stem (35b) are also few. Only two examples have been recorded, and their syllabic structure is particular in that the penultimate syllable has a long vowel. Therefore, these two patterns are somehow exceptional. In contrast with these two, words of HLLH stem are many. (35c) is not an exhaustive list.

The aforementioned observations show that the LLLH pattern and the HLLH pattern may be grouped together (both are final-H). If plurality is the criterion for judgment, the HLLH pattern may be considered basic, and the LLLH pattern should be its variation because this latter pattern has much fewer examples than the former pattern. With respect to three-syllable stem patterns, we said that the HLH pattern might be a realization of LLH, with the LLH pattern being the basic. We also inferred that the stem-initial (or antepenultimate) syllable of the word is realized as H-toned when the final syllable is also H-toned for some reason. If this inference is correct, the LLLH pattern might be basic despite having fewer examples than the HLLH pattern. It may well be that with a four- or more syllable stem, the final H tone tends to set off another H tone on the left.

5.4. Restricted tone system

Although Rukiga displays various tone patterns in nominal forms, its system is not unrestricted. "Restricted" and "unrestricted" are Voorhoeve's (1973) terms. In Rukiga, tone patterns do not increase in geometric progression as 2^n, where 2 is the number of basic tones (H and L), and *n* is the number of syllables in a word. Although one-syllable stem nouns have two patterns, H and L, two-syllable stem nouns do not have four ($=2^2$) patterns but three patterns, namely, LL, LH, and HL, without the HH pattern. For three-syllable stem nouns, four patterns, LLL, HLH, LHL, and HLL, of the eight ($=2^3$) patterns, are realized. For four-syllable stem nouns, eight patterns out of sixteen ($=2^4$) are realized, and eight patterns are lacking. For five-syllable stem nouns, five patterns of the thirty-two ($=2^5$) are realized.

Admitting that Rukiga's tone system is restricted, what type of restricted system does Rukiga have? It looks like the Runyankore type. In Runyankore, the number of patterns increases arithmetically, that is, $n+1$, where *n* is the number of stem syllables in a word. (36) is a list of tonal patterns of nominals in Runyankore.[23] Words of pattern *a* are all L-toned (or toneless). Words of pattern *b* are final-H. Words of pattern *c* are penultimate-H. Words of pattern *d* are antepenultimate-H. Words of pattern *e* are preantepenultimate-H. Words of pattern *f* are prepreantepenultimate-H. We can refer to patterns *a*, *b*, *c*, *d*, *e*, and *f* as patterns H0, H1, H2, H3, H4, and H5, respectively in consideration of the underlying location of H from the last syllable of the word. The pronunciation provided is that used in isolation, and the underlying H-toned syllable is underlined. The hyphen indicates the boundary of the nominal prefix and the stem. Notably, only one H tone, if any, appears in a word in isolation.

(36) 1. one syllable stem words
  a. H0 omu-zi 3,4 "root"

[23] The list is from Kaji (2010) with minor adjustments and additional data.

b. H1 omú-si 3,4 “vein, nerve”

2. two-syllable stem words
   a. H0 omu-hara 1,2 “daughter”
   b. H1 ama-rírа 6 “mourning”
   c. H2 omu-káma 1,2 “king”
3. three-syllable stem words
   a. H0 aka-gobora 12,14 “elephant tusk”
   b. H1 eki-tentére 7,8 “young hen”
   c. H2 oru-tongána 11,10 “index finger”
   d. H3 aka-tádoba 12,14 “hand-made lamp”
4. four-syllable stem words
   a. H0 aka-hungabebe 12,14 “termite”
   b. H1 oru-to:nyeréra 11,10 “drizzle”
   c. H2 aka-samunyíga 12,14 “skunk”
   d. H3 eki-gungúniro 7,8 “threshed corncob”
   e. H4 eki-sígisiro 7,8 “small clay pot”
5. five-syllable stem words
   a. H0 eky-a:rikirizo 7,8 “inner mat of saucepan”
   b. H1 eny-abarashaná 9,10 “species of weed”
   c. H2 VCV-CVCVCVCV́CV no examples
   d. H3 orw-a:kanákana “dewlap of cow”
   e. H4 VCV-CVCV́CVCVCV no examples
   f. H5 en-tá:gurukane 9,10 “crossroads”

Rukiga seemingly has a different tone system from Runyankore; some Rukiga words have two H tones in a word, but Runyankore words have at most one H tone in a word underlyingly. However, as discussed in Section 5.3, if words of two H tones are analyzed as having only one H tone, the Rukiga system is comparable to the Runyankore system, as in (37), where words with two Hs are classified as if the left H is irrelevant. (37) shows that the number of tone patterns is *n*+1, where *n* is the number of stem syllables in a word. +1 means that there are words that have no H tone, that is, words of H0 pattern. If a word has an H, it may be on any syllable of the stem, H1, H2, H3, H4, etc.

(37) a. one-syllable stem nouns
   H0 omu-zi 3, emi-zi 4 “root”
   H1 omú-ti 3, emí-ti 4 “tree”
   b. two-syllable stem nouns
   H0 oku-guru 15, ama-guru 6 “leg”
   H1 eki-sháka 7, ebi-sháka 8 “bush”

H2	omú-tima 3, emí-tima 4	"heart"
c. three-syllable stem nouns		
H0	oru-shahuzo 11, en-shahuzo 10	"curved machete"
H1	no example but:	
	? aka-gyégyéna 12, obu-gyégyéna 14	"strainer"
H2	eki-tágata 7, ebi-tágata 8	"hot spring"
H3	omu-nyúruru 3, emi-nyúruru 4	"door handle"
d. four-syllable stem nouns		
H0	eki-huruguto 7, ebi-huruguto 8	"earwax"
H1	e-kerezíya 9,10	"Catholic church"
	? ny-irá:nsáse 9,10	"variety of sweet potato"
	? omu-kóroró:mbya 3, emi-kóroró:mbya 4	"rainbow"
H2	en-tangahû:zi 9,10	"ginger"
	? omu-kárabâ:nda 3, emi-kárabâ:nda 4	"wooden case"
H3	eki-korógoto 7, ebi-korógoto 8	"dried papyrus plant"
H4	omu-cúraguzi 1, aba-cúraguzi 2	"night dancer"
e. five-syllable stem nouns		
H0	musingiriro 9,10	"variety of bean"
H1	kabaragára 9,10	"small sweet banana"
H2	no example	
H3	pa:sikárama 9,10	"lawn"
	? en-dábarásanyi 9,10	"hairy beggar-ticks"
H4	omw-é:yayamuro 3, emy-é:yayamuro 4	"yawn"

Although Rukiga and Runyankore have various tone patterns, their system is not unrestricted. One reason for this is Meeussen's rule whereby the right H tone of two adjacent H tones in a word is eliminated; therefore, the HH pattern does not exist in them. For two-syllable stem words, for example, there is no HH pattern in Rukiga and Runyankore because HH has become HL, merging with the original HL. (38) gives the correspondence of PB (Proto-Bantu) to Rukiga, where the right H tone is eliminated by its left H tone.

(38) a. PB *-kádí HH > Rukiga -kázi HL
b. omú-kazi 1, abá-kazi 2 "woman, wife"
c. omukázi wangye "my wife"

6. Southern dialect, northern dialect, and Runyankore

Generally, the southern dialect of Rukiga is considered more genuine and archaic than the northern dialect. The northern dialect of Rukiga exhibits a number of similarities

with Runyankore in pronunciation, vocabulary, and grammar. In this section, we briefly summarize the tonal characteristics of the northern dialect of Rukiga, together with Runyankore data. Four points are notable.

The first point is that HTA applies to the last syllable of words in isolation in the southern and northern dialects of Rukiga, like in Runyankore, as in (39a). However, only in the southern dialect does it apply to the penultimate syllable of words too. See (40a), in which the H of the penultimate **rí** of **orurímí** "tongue" is anticipated onto **ru** in isolation in the southern dialect but not in the northern dialect and Runyankore.

(39)		southern dialect	northern dialect	Runyankore	
	a.	omútwe 3	omútwe 3	omútwe 3	"head"
	b.	omutwé gwangye	omutwé gwangye	omutwé gwangye	"my head"
	c.	omutwé gúmwe	omutwé gúmwe	omutwe gúmwe	"one head"
(40)	a.	orúrimi 11	orurímí 11	orurímí 11	"tongue"
	b.	orurímí rwangye	orurímí rwangye	orurímí rwangye	"my tongue"
	c.	orurímí rúmwe	orurímí rúmwe	orurimi rúmwe	"one tongue"

The second point is that the High Tone Deletion (HTD) that works in Runyankore does not work in the two dialects of Rukiga. In (39c), the underlying H tone of **omutwé** "head" is deleted when the numeral adjective **gúmwe** "one" follows it in Runyankore, but not in Rukiga. The same can be said of (40c).

The third point is that in Runyankore, when all L nouns are followed by the possessive adjective "my" in a phrase, the last syllable of the noun is raised H, but this does not occur in the two dialects of Rukiga. See (41b), in which the last syllable of **okuguru** "leg" is raised H when followed by **kwangye** "my" in Runyankore.

(41)		southern dialect	northern dialect	Runyankore	
	a.	okuguru 15	okuguru 15	okuguru 15	"leg"
	b.	okuguru kwangye	okuguru kwangye	okugurú kwangye	"my leg"

The fourth point is that the northern dialect of Rukiga has two H-toned syllables in a word, as does the southern dialect, which is not the case in Runyankore (42). However, some northern dialect words have only one H-toned syllable in a word, like in Runyankore, and the southern dialect has two H-toned syllables in the word, as in (43).

(42)		southern dialect	northern dialect	Runyankore	
	a.	ekisírí:ra 7,8	ekisírí:ra 7,8	?	"burning firewood"
	b.	omubírikíra 3,4	omubírikíra 3,4	?	"funnel"

(43)	a.	ekánísa 9,10	ekanísa 9,10	ekanísa 9,10	"Protestant church"
	b.	aká꞉tíka 12,14	aka꞉tíka 12,14	aka꞉tíka 12,14	"broken piece"

The northern dialect of Rukiga is located between the southern dialect of Rukiga and Runyankore and shares a number of features, lexical, grammatical, and phonological (e.g., tonal) with Runyankore. Nevertheless, it is Rukiga in its tone rules, which crucially distinguish Rukiga from Runyankore.

7. Verb infinitives

Because verb infinitives are gerunds and included in class 15 of the noun classification in Rukiga, we provide in (44) a list of verb infinitives according to the syllable structure of the radical. Radicals are either L- or H-toned, regardless of the length and syllable structure of the radical of the word, except one. The exception is the verb **okuza/okúza** "to go somewhere," whose radical tone fluctuates either L or H. The derivation of **okúfa** "to die" is **o-ku-fú-a** > **okufwá** > **okufá** > **okúfa**.

(44)	L	-gu-	okugwa	"to fall"
		-ak-	okwaka	"to burn (intr.)"
		-ru-	okurwana	"to fight"
		-rim-	okurima	"to dig"
		-gond-	okugondama	"to bend down"
		-shob-	okushobo꞉rora	"to disentangle"
	H	-fú-	okúfa	"to die"
		-át-	okwâ꞉ta	"to break"
		-bí-	okubyâ꞉ma	"to sleep"
		-góy-	okúgoya	"to knead (stiff porridge)"
		-shék-	okushékura	"to pound"
		-tém-	okutémagura	"to cut in pieces"
	L/H	-gi-/-gí-	okuza/okúza	"to go somewhere"

(45) presents some minimal pairs of verb infinitives by tone. Conjugated forms also produce minimal pairs by tone, which are excluded here.

(45)	a.	L	okusa	"to grind"
		H	okúsa	"to get burned"
	b.	L	okushara	"to go mad"
		H	okúshara	"to cut"
	c.	L	okurenga	"to weigh (tr.)"
		H	okurê꞉nga	"to fall from the edge"

d. L okutarika “to smoke (meat)”
H okutárika “to flower [of crops]”

8. Conclusion

We overviewed the tonal patterns of nominals in Rukiga. Rukiga apparently has more tone patterns in nominals than its closely related Runyankore because it, unlike Runyankore, has nouns with two H tones in a word in isolation. However, by examining its tonal patterns in terms of complementary distribution and considering other factors, we found that some two H tone patterns are considered to have only one H tone underlyingly, and the system can be reduced to the system like that of Runyankore such that nominals have only one, if any, H-toned syllable in a word. The analysis at present is not definite; additional observations are necessary to further understand the patterns.

Concerning HTA, notably, the H tone of not only the final but also the penultimate syllables of words moves to the left by one syllable in isolation in the southern dialect of Rukiga. This is not, or seldom, observed in other languages of the area, according to our knowledge.[24]

Excursus: About Kisseberth and Ndabasara (1993)

Kisseberth and Ndabasara (1993) is a tonological study of Rukiga, according to the authors. However, their language, written as Ruciga, seems more like Runyankore than Rukiga. They do not specify the location (or ethnic group) where the language is spoken. There are several reasons why the present author thinks it is more like Runyankore than Rukiga. We present three examples: (46), (47), and (48).

(46) HTA does not apply to H2 tone pattern words
a. omu-kóno 3,4 “arm”
omu-kóno gwaanje 3 “my arm”
b. omu-kóno 3,4 “arm”
omu-kóno gwangye 3 “my arm”
c. omú-kono 3,4 “arm”
omu-kóno gwangye[25] 3 “my arm”

[24] In some dialects of Runyoro, spoken approximately 500 km north of Rukiga, this occurs. For example, the Nyoro language is called **orunyôro** 11 in the Hoima area, but the pronunciation **orúnyoro** 11 is heard in the Masindi area. Moreover, we need to verify with Rufumbira, spoken southwest of Rukiga. It is a dialect of Kinyarwanda but seems to have a number of common features with Rukiga.

[25] The vowel **a** in **gwangye** is long, but as it is not fully two-mora long (not [gwaːnje] but [gwaˑnje]), the length mark ː is not used, as stated in Note 5. The same is true of **a** before **nd** in **omuganda** 3,4“bundle” in (47c).

(46a) is from Kisseberth and Ndabasara (1993). (46b) is Runyankore data of the present author, and (46c) is Rukiga data of the present author. In Kisseberth and Ndabasara's (1993) data, HTA does not apply to the H tone of the penultimate syllable of words in isolation, which is a characteristic of Runyankore, verified in (46b), and also with pattern *c* words in (36). In Rukiga, HTA applies in such a case, as stated in Section 5.1.1, and observed in (46c).

The second reason is provided in (47). (47a) is from Kisseberth and Ndabasara (1993), in which the final L tone of **omu-gaanda** 3,4 "bundle" is raised H when the L-toned adjective **mu-reingwa** "long" follows it. This is a characteristic of Runyankore, verified in (47b), which is the present author's description of the Runyankore data. (47c) is the Rukiga data, where High Tone Insertion (HTI) does not work.[26]

(47) HTI applies to H0 tone pattern words

a.	omu-gaanda 3,4	"bundle"
	omu-gaandá mu-reingwa 3	"long bundle"
b.	omu-gaːnda 3,4	"bundle"
	omu-gaːndá mu-riːngwa 3	"long bundle"
c.	omu-ganda 3,4	"bundle"
	omu-ganda mu-rengwa 3	"long bundle"

The third reason is exemplified in (48). In Runyankore, High Tone Effacement (HTE) applies to nominals with an H tone when an adjective with an H tone follows it. (48a) is from Kisseberth and Ndabasara (1993), in which the H tone of **en-kóko** 9,10 "chicken" is effaced because the adjective **mpáango** "big" with an H element follows it. This is a characteristic of Runyankore, verified in (48b), which is the present author's description of the Runyankore data. (48c) is the corresponding Rukiga data, in which HTE does not apply.

(48) HTE applies to words with H tone

a.	en-kóko 9,10	"chicken"
	en-koko mpáango 9	"big chicken"
b.	en-kóko 9,10	"chicken"
	en-koko mpâːngo 9	"big chicken"
c.	en-kóko 9,10	"chicken"
	en-kóko mpâːngu 9	"big chicken"

[26] The word **omuganda** 3,4 "bundle" is known but rarely used in Rukiga.

Acknowledgements

This is a product of research activity of the Institute for Comprehensive Research, Center for Language Studies, which was financially supported by the Kyoto Sangyo University Research Grant K2001. The travel expenses were funded by JSPS KAKENHI, Grant numbers 20K00556 and 24K03902.

References

Hyman, Larry M. 2022. High Tone Bumping in Runyankore. *Journal of African Languages and Linguistics* 43-1: 37-81.

Kaji, Shigeki. 2004. *A Runyankore Vocabulary.* Tokyo: Research Institute for Languages and Cultures of Asia and Africa.

Kaji, Shigeki. 2010. A Comparative Study of Tone of West Ugandan Bantu Languages, with Particular Focus on the Tone Loss in Tooro. *ZAS Papers in Linguistics* 53: 99-107.

Kaji, Shigeki. 2015. *A Runyoro Vocabulary*. Kyoto: Shoukadoh.

Kaji, Shigeki. 2023. *A Rukiga Vocabulary.* Kyoto: Shoukadoh for the Center for Language Studies, Kyoto Sangyo University.

Kaji, Shigeki. 2024. Palatalization and Velarization in Rukiga Pronunciation. *Working Papers in African Linguistics* 2: 1-28.

Kisseberth, Charles W. and Emmanuel Ndabasara. 1993. The Nominal Tonology of Ruciga. Champaign-Urbana: Ms. University of Illinois.

Lewis, M. Paul, Simons, Gary F. and Charles D. Fennig (eds.) 2015. *Ethnologue: Languages of the World*. 18th edition. Dallas, Texas: SIL International.

Voorhoeve, Jan. 1973. Safwa as a Restricted Tone System. *Studies in African Linguistics* 4: 1-21.

Between antipassive and differential object marking in Bari: A markedness paradox and the antipassive-to-active shift

Shuichiro Nakao
Osaka University

Abstract
This study describes Bari, an Eastern Nilotic language spoken in South Sudan, as having a unique morphosyntactic alternation representing a historical/typological pivot between antipassive (AP) and differential object marking (DOM). This alternation is affected by the topicality of the object, as well as tense-aspect-modality, clause types, etc. (like both DOM and AP), and the marked transitive stems (with the suffix **-jV**) code non-topic or non-referential objects (like AP, unlike DOM), but the object is not demoted (like DOM, unlike AP). This distinction is attested with various clause types, such as preterit, perfect and imperative, but not with the others (imperfective, non-past, etc.). Comparative evidence suggests that it is derived from a proto-Southern Eastern Sudanic antipassive construction, and this change can be characterized as a case of antipassive-to-active shift, whereby the antipassive is reanalyzed as the general active form. While this phenomenon is not uncommon in ergative-to-accusative alignment change, Bari exhibits the unique case in that it took place within an accusative language.

Keywords: antipassive, differential object marking, markedness, Nilo-Saharan

1. Introduction

Bari is an Eastern Nilotic (among Nilo-Saharan) language spoken in South Sudan, closely related to the Pojulu, Nyangbara, Nyepu, Kuku, Kakwa and Mandari languages, which constitute the Bari group (Vossen 1982, the neutral term 'Karo' has also been proposed). The Bari group is characterized by S/A-V(-O) word order (except for in relative clauses, see Section 3.3) and the lack of case marking on nouns, in contrast with the rest of the Eastern Nilotic branch (the Teso-Turkana-Lotuho-Maa(sai) group), which exhibit general V-S/A(-O) order and marked nominative case marking (König 2008). Although the linguistic study of Bari has a relatively long history, having started in the mid-19th century (see Spagnolo 1933: xix–xxii), its morphosyntax has remained poorly investigated.[1]

[1] The Bari and Juba Arabic data were provided by Dr. David Gore, a native speaker of both languages, during the period August 2020–March 2022 in Osaka, Japan. This study uses an IPA-based phonemic transcription for Bari except that low tone is not marked. The [±ATR] vowels are

This study aims at describing Bari as having a unique morphosyntactic alternation representing a historical and typological pivot between antipassive (abbreviated: AP) and differential object marking (abbreviated: DOM).

In linguistic typology, AP refers to the de-transitivizing alternation whereby the object of a transitive (active) clause is deleted or demoted with an oblique case to code indefiniteness and/or non-topicality and, as a side effect, the clause tends to have imperfective reading (Dixon 1994; Polinsky 2017; Janic & Witzlack-Makarevich 2021), while DOM refers to a morphosyntactic phenomenon whereby an overt marking (accusative case and/or object agreement) is realized only with objects with higher animacy, definiteness and/or toplicality, or with certain tense-aspect-modality (abbreviated: TAM) or clause types (Aissen 2003; Dalrymple & Nikolaeva 2011; Sinnemäki 2014; Witzlack-Makarevich & Seržant 2018).[2]

While DOM has been widely discussed as a common phenomenon among accusative languages (Malchukov 2017), AP has until recently been thought of as an 'exotic' phenomenon most often attributed to ergative languages (Polinsky 2017). A few recent typological studies, however, indicate the resemblance of the two constructions in that both of them not only code the object in different markings, but also share similar triggers, such as definiteness, information structure, TAM and clause type. Meanwhile, the two constructions are distinguished in terms of markedness; namely, DOM marks definite/topic objects, whereas AP marks indefinite/non-topic/demoted objects if any (Schikowski & Iemmolo 2015; Polinsky 2017; Janic & Witzlack-Makarevich 2021).

In the Nilo-Saharan linguistics, Bari has been analyzed as having an 'antipassive' construction (e.g., Schröder 2015; Payne 2021), as exemplified in (1), adapted from Spagnolo (1933: 137, checked and edited by me and my consultant).[1] The third person pronoun (object) is zero-coded in certain grammatical contexts, including (1c).[2] Hereafter,

written as follows: ɪ, ɛ, a, ɔ, ʊ = [−ATR], i, e, ʌ, o, u = [+ATR]. The Bari orthography differs from our transcription in the use of the letters j [ɟ], ny [ɲ], y [j], 'b [ɓ], 'd [ɗ], 'y [ʔj], (' [ʔ]), i [ɪ, i], e [ɛ, e], ö [ʌ], o [ɔ, o], u [ʊ, u] and the lack of tone marking. Our consultant consistently lacks the word-final glottal stop documented in the other Bari sources (Spagnolo 1933; Vossen 1982; Yokwe 1987). The glosses follow the Leipzig Glossing Rules.

[2] Following Dalrymple & Nikolaeva (2011), we use 'DOM' for both the head- and dependent-marking systems. Schikowski & Iemmolo (2015) propose a separate term 'differential object indexing' for the head-marked counterpart and uses 'differential object coding' as a cover term.

[1] Earlier studies labelled this suffix in various ways, e.g., 'emphatic' (Spagnolo 1933), 'qualitative' (Tucker & Bryan 1966, an old term for 'antipassive' in Nilo-Saharan/Sudanese linguistics) and 'indefinite' (Yokwe 1978). Cohen (2000: 82–88) labels the cognate form in Kuku as a 'telic' marker, but this analysis is not tenable for Bari.

[2] According to our consultant, human nouns may be referred to with a personal pronoun, as contrasted by **nân a mɛ́t Ø.** (1SG PRT see 3) 'I saw it/them (things).' vs. **nân a mɛ́t ɲe/sɛ.** (1SG PRT see 3SG/3PL) 'I saw him/them (human).' Bari lacks gender distinction in personal pronouns and hereafter translated with 'he/him/his'.

the verbal suffix in (1b,d) will be referred to as '**-ɟV**' (gloss: SUF).

(1) a. nân a **dɛ́r** súkuri[3].
1SG PRT **cook** chicken
'I cooked the chicken (the one under discussion).'
b. nân a **dɛ́r-ɟá** súkuri.
1SG PRT **cook-SUF** chicken
'I cooked a chicken.'
c. nân a **dɛ́r** Ø.
1SG PRT **cook** 3
'I cooked it (i.e., something already referred to).'
d. nân a **dɛ́r-ɟa**.
1SG PRT **cook-SUF**
'I cooked (i.e., I did some cooking.)'

Spagnolo (1933: 137) correctly notes that this suffix is used 'when there is no object discussed' or the 'object has not already been discussed in a previous sentence' or 'where an object is general rather than particular, or where an action is habitual rather than exclusive.' To complement his words, this suffix also codes anticausative, reciprocal, nominalization, focus and some TAMs (the 'habitual' reading mentioned by him requires precaution; see Section 4), as discussed in the next sections.

The traditional label 'antipassive' for this suffix would be puzzling to typologists. Example (1d) seems to be a *bona fide* instance of antipassive, and the functional syncretism noted by Spagnolo is not uncommon with the antipassive (Sansò 2017; Polinsky 2017; Janic & Witzlack-Makarevich 2021). However, the example (1b) deviates from the basic criterion of antipassive, since there is no evidence suggesting that the object has been 'demoted.' The object lacks the oblique marking, and indeed the discourse-new or focused obligatory arguments are introduced by a verb with this suffix (see Sections 3 and 4).[4]

From another viewpoint, we could rather compare the non-suffixed form with the head-marked DOM (or 'object agreement') in languages such as Swahili, whereby an (in-situ) discourse-old topic-like object is coded with an object pronominal affix (2).

[3] The underlying form of this word is **súkúri**. Bari has a general tone sandhi rule whereby the word-final H causes the tonal change of the following word: L → H/H#, H→ L/H#_L, H/F → L/H#H_, H/L → F/H#_ and F → L/H#_ (see also Yokwe 1987). This rule is not applied to certain constructions (e.g., between subject and predicate).

[4] 'Antipassive' constructions are often deployed to mark indefiniteness or discourse-new objects in Western Nilotic languages (Andersen 2018).

(2) Swahili (Ashton 1944: 44–45, glosses are my own)

a. u-me-leta chakula?
   2SG-PERF-bring food(CL.7)
   'Have you brought (the) food?'

b. ni-me-leta bwana
   2SG-PERF-bring sir
   'I have brought (it), sir.'

c. u-me-**ki**-leta chakula?
   2SG-PERF-**CL.7**-bring food(CL.7)
   'Have you brought the food (which I asked you to bring)?'

d. ni-me-**ki**-leta bwana
   1SG-PERF-**CL.7**-bring sir
   'I have brought it, sir.'

Although the functions coded by the non-suffixed form in Bari (1a, c) and the object markers in Swahili (2c, d) seem nearly identical, they clearly differ in the directionality of morphological marking: Bari deviates from the basic criterion of DOM in that it is the non-suffixed form that codes the marked function. This markedness paradox is crucial, since canonical DOM has been characterized by iconicity; that is, only the less typical objects (e.g., animate, definite) are morphologically marked (Aissen 2003).[5] In view of the emerging number of typological studies that are aware of the similarity between DOM and AP (Schikowski & Iemmolo 2015; Polinsky 2017; Janic & Witzlack-Makarevich 2021), such a pivotal case as the Bari verbal alternation deserves attention.

The next sections describe the basic verbal morphosyntax of Bari including the voice alternation (Section 2), the triggers for this alternation (Section 3) and its neutralization in some clause types (Section 4). In Section 5, we present the diachronic outlook by comparing the cognate suffixes in other Nilo-Saharan branches and argue that Bari went through an antipassive-to-active shift whereby the antipassive became the default alignment, a process attested in the ergative-to-accusative shift (see Dixon 1994; Harris & Campbell 1995; Aldridge 2011; Hemmings 2021).

2. Basic verbal morphosyntax

2.1 Basic verbal paradigms

Bari exhibits a complex system of verbal inflection. Basic (underived) transitive verbs are categorized into H vs. LHL (in terms of underlying tone)[6] and monosyllabic vs.

[5] Note, however, that Bárány (2012) analyses the 'definite' vs. 'indefinite' forms of transitive verbs in Hungarian, where both forms have an overt marking, as an instance of DOM.

[6] We follow Yokwe (1987) for these labels. The distinction is most clearly exhibited in the

multisyllabic roots according to the underlying phonological form (intransitive verbs are discussed later). There is a three-way voice-like alternation (**-Ø** vs. **-ɟV** vs. **-V**), whose functions are discussed in 2.2. The vowel pattern of these suffixes is largely (but not completely) predictable with [±ATR] (and [±high]) vowel harmony (see Spagnolo 1933: 141). The onset consonant of the **-ɟV** suffix is assimilated to the root-final consonant except after a vowel or /r/, as in Table 1.[7] There are a few irregular verbs such as **ɲʌ́** 'eat' that take unique suffixes (namely **-su**, see Section 5).[8]

Table 1: Verbal suffixes (the 'base' form)

Ø	SUF	PASS		Ø	SUF	PASS	
dɛ́r	dɛ́r-**ɟa**	dɛ́r-**â**	'cook'	kín	kɪn-**dô**	kɪn-**ô**	'close'
kʊ́r	kʊr-**ɟô**	kʊr-**ô**	'cultivate'	ɟɔ́ŋ	ɟɔŋ-**gâ**	ɟɔŋ-**â**	'take'
pí	pɪ-**ɟâ**	pɪ-**jâ**	'ask'	lʊ́ŋ	lʊŋ-**gô**	lʊŋ-**ô**	'call'
jé	jé-**ɟu**	jé-**jî**	'think'	dɛ́k	dɛ́g-**ga**	dɛ́k-**â**	'talk'
ɓɛ́(ʔ)	ɓɛ-**ʾja**	ɓɛ́l-**â**	'split'	jɛ́m	jɛ́m-**ba**	jɛ́m-**â**	'marry'
pé(ʔ)	pé-**ʾju**	pél-**ô**	'burn'	ɟám	ɟám-**bʊ**	ɟám-**â**	'talk'
dɛ́n	dɛ́n-**dja**	dɛ́n-**â**	'know'	dɛ̃́p	dɛ̃́b-**ba**	dɛ̃́p-**â**	'grasp'
mɛ́t	mɛd-**djâ**	mɛt-**â**	'meet'	ɲʌ́	ɲé-**su**	ɲʌ́-**jî**	'eat'

The verbs are inflected according to the TAM and information/clause structure, as summarized in Tables 2 (**dɛ́r** 'cook,' **mɛ́t** 'see'), 3 (**ɓélɛ́ŋ** 'break,' **kʌbûr** 'mix') and 4 (intransitive verbs).

The base forms always follow a TAM/polarity particle (e.g., **a** 'PRT', **aɟé** 'PERF', **tɪ** 'NEG', **ákɔ** 'PRT.NEG') when it functions as a finite verb, but they also function as a non-finite form (verbal noun, relative clause).

Affirmative non-past and imperfective TAMs and focus are coded either by the non-past form, which also functions as 'disjoint' imperfective (NPST/DJ:IPFV, marked by reduplication of the word-initial mora) and the focus form, which also functions as 'conjoint' imperfective (FOC/CJ:IPFV, marked by the word-initial low tone). These forms will be discussed in Section 4.

benefactive (a type of applicative), as in H **dɛ́r-ákín** 'cook for (someone)' vs. LHL **mɛt-ákɪn** 'see for (someone)'. We do not discuss derived (including applicative) verbs in this study.

[7] As our consultant consistently lacks the word-final glottal stop attested in the other works (deriving from underlying /l/), it is bracketed in these examples (**ɓɛ́(ʔ)** and **pé(ʔ)**).

[8] Some loan verbs have **-ʊ** instead of **-ɟV**, e.g., **malâ** (Ø) vs. **malá-ʊ** (SUF) vs. **malá-jí** (PASS) 'fill' (from Juba Arabic **mála(u)**). The same ending is found in an intransitive verb **naɟáʊ** 'succeed' (from Juba Arabic **náɟa**) although it lacks the morphological voice alternation. Other verbs conform to the ordinary pattern (e.g., **karíɟʊ** (Ø) vs. **karíɟʊ-ɟa** (SUF) vs. **karíɟʊ-wá** (PASS) 'graduate' from Juba Arabic **káriɟu**).

Table 2: Monosyllabic transitive verb classes (+OBJ = when followed by an object)

class	voice	base	CJ:IPFV / FOC	DJ:IPFV / NPST	IMP
H	Ø	**dɛ́r**	**dɛr**	**dɛ~dɛ́r**	**dɛr-ɛ́**
	SUF +OBJ	**dɛ́r-ɟa** **dɛ́r-ɟá**	**dɛr-ɟá**	**dɛ́~dɛr-ɟa** **dɛ́~dɛ́r-ɟá**	**dɛr-ɟ-í**
	PASS	**dɛ́r-â**	**dɛr-á/-a**	**dɛ́~dɛ́r-a**	**dɛr-a-ní**
FREQ	Ø	**dɛ~dɛ̂r**	**dɛ~dɛr**		**dɛ~dɛr-ɛ́**
	SUF	**dɛ~dɛ́r-ɟa**	**dɛ~dɛ́r-ɟa**		**dɛ~dɛr-ɟ-í**
	PASS	**dɛ~dɛ́r-a**	**dɛ~dɛ́r-a**		**dɛ~dɛr-a-ní**
LHL	Ø	**mɛ́t**	**mɛt**	**mɛ~mɛ́t**	**mɛt-ɛ́**
	SUF	**mɛd-djâ**	**mɛd-djâ**	**mɛ́~mɛd-djâ**	**mɛd-d-í**
	PASS	**mɛt-â**	**mɛt-a**	**mɛ́~mɛt-â**	**mɛt-a-ní**
FREQ	Ø	**mɛ~mɛ̂t**	**mɛ~mɛt**		**mɛ~mɛt-ɛ́**
	SUF	**mɛ~mɛ́d-dja**	**mɛ~mɛ́d-dja**		**mɛ~mɛd-d-í**
	PASS	**mɛ~mɛ́t-a**	**mɛ~mɛ́t-a**		**mɛ~mɛt-a-ní**

Table 3: Multisyllabic transitive verb paradigm (the use of the '+OBJ' form is optional)

class	voice	base	CJ:IPFV / FOC	DJ:IPFV / NPST	IMP
H	Ø +OBJ	**ɓélɛŋ**	**ɓelɛŋ**	**ɓé~ɓélɛŋ** **ɓé~ɓélɛ́ŋ**	**ɓelɛŋ-é**
	SUF +OBJ	**ɓélɛ́ŋ-gû** **ɓélɛ́ŋ-gú**	**ɓelɛ́ŋ-gu**	**ɓé~ɓélɛŋ-gû** **ɓé~ɓélɛ́ŋ-gú**	**ɓelɛŋ-gí**
	PASS	**ɓélɛ́ŋ-ó**	**ɓelɛ́ŋ-o**	**ɓé~ɓélɛŋ-ó**	**ɓelɛŋ-o-ní**
FREQ	Ø	**ɓe~ɓélɛŋ**	**ɓe~ɓelɛŋ**		**ɓe~ɓelɛŋ-é**
	SUF	**ɓe~ɓélɛŋ-gu**	**ɓe~ɓélɛŋ-gu**		**ɓe~ɓelɛŋ-gí**
	PASS	**ɓe~ɓélɛŋ-o**	**ɓe~ɓélɛŋ-o**		**ɓe~ɓelɛŋ-o-ní**
LHL	Ø +OBJ	**kʌbûr**	**kʌbur**	**kʌ́~kʌbûr** **kʌ́~kʌ́búr**	**kʌbur-é**
	SUF +OBJ	**kʌbúr-ɟʌ** **kʌ́búr-ɟʌ́**	**kʌbúr-ɟʌ**	**kʌ́~kʌbúr-ɟʌ** **kʌ́~kʌ́búr-ɟʌ́**	**kʌbur-ɟí**
	PASS	**kʌbúr-ʌ́**	**kʌbúr-ʌ**	**kʌ́~kʌbúr-ʌ́**	**kʌbur-ʌ-ní**
FREQ	Ø	**kʌ~kʌ́bur**	**kʌ~kʌbur**		**kʌ~kʌbur-é**
	SUF	**kʌ~kʌ́bur-ɟʌ**	**kʌ~kʌ́bur- ɟʌ**		**kʌ~kʌbur-ɟ-í**
	PASS	**kʌ~kʌ́bur-ʌ**	**kʌ~kʌ́bur-ʌ**		**kʌ~kʌbur-ʌ-ní**

The imperative (IMP) form, which also functions as 'sequential,' is marked by the low-toned stem and the high-toned imperative suffix. Some other TAM expressions are derived from this form (Nakao, forthcoming).

In addition, each verb has the frequentative stem (FREQ, pluractional/repetitive, i.e.,

'to do many times/repeatedly/frequently/habitually'), which is formed by reduplication and always represents an imperfective TAM distinguished from the non-past form only by tone. The frequentative has the same tonal forms for both H and LHL roots, and the 'conjoint' vs. 'disjoint' distinction is neutralized.

The high tone spread before the object (e.g., **dérɟa** vs. **dérɟá** + OBJ in Table 2, which also triggers tone sandhi), as contrasted in (1b) vs. (1d), represents a phrasal tone rule, which is also applied in a few cases of noun phrase formation, as exemplified in (3).[9] As for the multisyllabic verb roots (Table 3), the use of the '+OBJ' form is optional.

(3) a. **mɔ́ŋɛ** vs. **mɔ́ŋɛ́** ɟáda
father father Jada (underlying: /ɟada/)
'(his) father' 'Jada's father'

b. **ŋína** vs. **ŋíná** súkuri
MED.F MED.F chicken (underlying: /súkúri/)
'this/that (one)' 'this/that chicken'

c. **kólɔn** vs. **kólɔ́n** ɟur-ʌn
which.M.PL which.M.PL country:M-PL (underlying: /ɟúr-ʌn/)
'which (one.M)' 'which country'

Intransitive verbs have distinct paradigms according to the inflectional classes as shown in Table 4. The first three classes overlap with some parts of the transitive verb paradigm.

Table 4: Intransitive verbal paradigms

class	base	CJ:IPFV / FOC	DJ:IPFV / NPST	IMP	
CV́C-V̂	**dót-ô**	**dot-ó/-o**	**dó~dót-o**	**dot-o-ní**	'sleep'
FREQ	**do~dót-o-tû**	**do~dót-o-tû**		**do~dot-o-ní**	
CVC-V̂	**wʌr-â**	**wʌr-ʌ**	**wʌ́~wʌr-â**	**wʌr-ʌ-ní**	'walk'
FREQ	**wʌ~wʌ́r-ʌ-tû**	**wʌ~wʌ́r-ʌ-tû**		**wʌ~wʌr-ʌ-ní**	
CV́C-ɟV	**ɗúr-ɟʌ**	**ɗur-ɟʌ́**	**ɗú~ɗur-ɟʌ**	**ɗur-ɟ-í**	'grow'
FREQ	**ɗu~ɗúr-ɟʌ-tû**	**ɗu~ɗúr-ɟʌ-tû**		**ɗu~ɗur-ɟ-í**	
CV́C-V́n	**rɔ́m-án**	**rɔm-án**	**rɔ́~rɔ́m-an**	**rɔm-ɛ́**	'greet'
FREQ	**rɔ~rɔ́m-ɪ**	**rɔ~rɔ́m-ɪ**		**rɔ~rɔm-ɛ́**	
CV̂C	**rûm**	**rum**	**rú~rûm**	**rum-é**	'meet'
FREQ	**ru~rúm-i**	**ru~rúm-i**		**ru~rum-é**	

[9] The general tone sandhi rule does not affect the tonal phrases mentioned here.

irregular	**tô**	**tʊ**	**tʊ́~tʊ**	**ití ~ (i)tiní**	‘go’
FREQ	**tʊ~tʊ́~tʊ-tû**	**tʊ~tʊ́~tʊ-tû**		**ti~ti-ní**	
irregular	**pɔ̂**	**pɔ́**	**pɔ́~pɔ̂**	**pɔ-ô**	‘come’
FREQ	**pɔ~pɔ́~pɔ-tû**	**pɔ~pɔ́~pɔ-tû**		**pɔ~pɔ-ô**	
irregular	**pɔ́n-djâ**[10]	**pɔ́n-djâ**	**pɔ́~pɔn-djâ**	**pɔn-d-í**	‘come’
FREQ	**pɔ~pɔ́n-dja**	**pɔ~pɔ́n-dja**		**pɔ~pɔn-d-í**	

2.2 Anticausative, passive, reflexive, reciprocal

As mentioned in Section 1, the functional syncretism of anticausative/reciprocal and AP (in parallel with the suffix **-ɟV**) is cross-linguistically attested (Sansò 2017; Polinsky 2017; Janic & Witzlack-Makarevich 2021). In Bari, anticausative signals the lack of agent (4a) contrasted with passive (4b), whereby the agent is implied. The demoted agent in a passive sentence is optionally introduced by the instrumental-cum-comitative preposition **kɔ** ‘with/by’. Except for some roots such as ‘cook’, the intransitive stem with **-ɟV** mostly codes either antipassive or anticausative, as contrasted in (5) vs. (6).[11]

(4) a. súkúri a **dɛ́r-ɟa**.
chicken PRT **cook-SUF**
‘The chicken is cooked (i.e., it is ready).’ (anticausative)
b. súkúri a **dɛ́r-â** (kɔ nân).
chicken PRT **cook-SUF** (with 1SG)
‘The chicken was cooked (by me).’ (passive)

(5) a. kotumît a **ŋá-ɟʊ**.
door PRT **open-SUF**
‘The door opened.’ (anticausative)
b. *nân a **ŋá-ɟʊ**.
1SG PRT **open-SUF**
intended: ‘I opened.’ (antipassive)

(6) a. *bûk a **kɛn-djâ**.
book PRT **read-SUF**
intended: ‘The book is read.’ (anticausative)
b. nân a **kɛn-djâ**.
1SG PRT **read-SUF**
‘I have read (i.e., I went to school).’ (antipassive)

[10] For now, **pɔ̂** and **pɔ́n-djâ** seem to be in free variation.
[11] This fact parallels the ‘middle’ voice in Southern Lwo (Noonan 1992; Hieda 2009).

While Spagnolo (1933: 81) explains the anticausative construction as deletion of the reflexive pronoun (SG. **mógón**, PL. **ɓʌ́rîk** < 'body'), this explanation is not justifiable since the reflexive pronoun can supplement only automatic, [−intentional] and [−affected] events (e.g., 'open', 'loosen'). Moreover, the imperfective counterparts (8) entail different morphological coding (see Section 4).

(7) a. kotumît a **ŋá-ɟó**[12] mógɔn.
door PRT **open-SUF** body
'The door opened (automatically/accidentally).' cf. (5a)
b. *súkúri a **dɛ́r-ɟá** mógɔn.
chicken PRT **cook-SUF** body
intended: 'The chicken is cooked.' cf. (4a)

(8) a. kotumît **ŋá~ŋa-ɟɔ**.
door **DJ.IPFV~open-SUF**
'The door is opening.' (anticausative)
b. kotumît **ŋa-ɟó** mógɔn.
door **CJ.IPFV:open-SUF** body
'The door is opening (automatically/accidentally).' (reflexive)

Although it is not fully productive, **-ɟV** can code reciprocity (9). In this case, too, the use of a reflexive pronoun seems to code different meanings (e.g., 'meet' vs. 'assemble'). In the reciprocal construction, the object (including reflexive/reciprocal pronouns) can be introduced by the comitative/instrumental preposition **kɔ** 'with'.

(9) a. jî a **mɔ́r** Ø
1PL PRT **gather** 3
'We gathered/assembled them (things).'
b. jî a **mɔ́r-ɟa**.
1PL PRT **gather-SUF**
'We met (got together).'
c. jî a **mɔ́r-ɟá** ɓʌ́rik.
1PL PRT **gather-SUF** body.PL
'We assembled/united.' (not interpreted as 'We met.')
d. jî a **mɔ́r-ɟa** kɔ {pɔní / ɓʌ́rîk}
1PL PRT **gather-SUF** with {Poni / body.PL}
'We met with {Poni / each other}.'

[12] Here, **-ɟV** is obligatory since its absence would code intentional action (Section 3).

2.3 Nominalization

The 'base' form suffixed with the **-ɟV** 'SUF' or **-V** 'PASS' also codes nominalization (both are treated as a feminine noun), with different functions. When the verb lacks an object, the passive form is used (10). With an object, both forms can be used and the **-ɟV** form is interpreted as anticausative, as contrasted in (11) and (12). The subject and object are marked by a genitive particle.[13]

(10) **{dɛ́r-â / *dɛ́r-ɟa }**[14] (ná mamá) a ná-ɓʊ́t.
{cook-PASS / cook-SUF} (POSS.SG.F my.mother) COP F-good.SG
'(My mother's) Cooking is good.'

(11) a. **dɛ́r-â** ná sukúri a ná-gó.
cook-PASS POSS.F.SG chicken(.F) COP F-hard.SG
'It is hard (for one) to cook chicken.'
b. **dɛ́r-ɟa** ná sukúri a ná-gó.
cook-SUF POSS.F.SG chicken(.F) COP F-hard.SG
'The chicken is hard to cook (it is hard for the chicken to be cooked).'

(12) a. nân a mɛd-djâ **dɛ́r-â** ná sukúri.
1SG PRT see-SUF **cook-PASS** POSS.SG.F chicken(.F)
'I saw the cooking of chicken (I saw someone cooking chicken).'
b. nân a mɛd-djâ **dɛ́r-ɟa** ná sukúri.
1SG PRT see-SUF **cook-SUF** POSS.SG.F chicken(.F)
'I saw the cooking of chicken (I saw the chicken cooked).'

In contrast, when the verb is a complement of an auxiliary verb (plus a preposition[15]), the **-ɟV** form is used, and the object is not marked by a genitive particle, as exemplified below. Note that the use of the third person pronoun is obligatory when the verb has **-ɟV** (13c,d).

[13] The genitive particle (SG.M **lɔ́**, SG.F **ná**, PL. **tí**) also affects the tone of the following word (e.g., the syllable after **lɔ́** and **ná** is low-toned). These particles are also affected by the general tone sandhi rule.
[14] Hereafter {A/(*)B} demonstrates different choices (and grammaticality) in the same context.
[15] Bari has allative **ɪ** and locative **í**, and each affects the tone of the following word (the syllable after **ɪ** is high-toned and after **í** low-toned). These prepositions are not affected by the tone sandhi rule.

(13) a. nân ɗeke **{dɛ́r-ɟá / *dɛ́r**[16] **/ *dɛ́r-â}** súkuri.
1SG CJ.IPFV:want **{cook-SUF / cook / cook-PASS}** chicken
'I want to cook (the) chicken.'

b. nân ɗeke **dɛ́r-ɟa**.
1SG CJ.IPFV:want **cook-SUF**
'I want to cook.'

c. nân ɗeke **dɛ́r-ɟá** ɲe.
1SG CJ.IPFV:want **cook-SUF** 3SG
'I want to cook it (thing).'

(14) a. nân a sɔ́lɔ́-ɟá **{dɛr-ɟa / *dɛ̂r}** súkúri.
1SG PRT start-SUF **{cook-SUF / cook}** chicken
'I started cooking (the) chicken.'

b. lɔpɛ́ŋ gwɔn í **{wur-ɟʌ / *wur-ʌ}** ná buk.
3SG CJ.IPFV:stay in **{write-SUF / write-PASS}** POSS.SG.F book
'He is still writing the book.'

c. nân tʊ ɪ **{dɛ́r-ɟa / *dɛ́r-â}** (ná sukúri).
1SG CJ.IPFV:go to **{cook-SUF / cook-PASS}** (POSS.SG.F chicken)
'I am going to cook (chicken).'

d. nân pɔ́~pɔ̂ ɪ **{dɔ́g-ga / *dɔ́k-â}** ná dɔ.
1SG NPST~come to **{pick-SUF / pick-PASS}** POSS.SG.F 2SG
'I will come to pick you up.'

The infinitive form of intransitive verbs is identical to the base form (e.g., **dót-ô** 'sleeping,' **wʌr-Â** 'walking,' **rɔ́m-án** 'greeting') except the /CV̂C/ type, which are nominalized by **-ɟV** (e.g., **rûm → rúm-bʌ** 'meeting,' **pûn → pún-djʌ** 'growing'). As for these verbs, this suffix never appears when they are predicates.

3. Triggers of the alternation

The alternation of **-ɟV** vs. **-Ø** has multiple triggers such as topicality, referentiality, clause type and intentionality, which are parallel to the DOM constructions (see Section 1).[17] Since certain clause types do not attest to this alternation, first we draw on the preterit construction (marked by the preverbal particle **a**) and examine how and where it is manifested and argue that the suffixed form is functionally unmarked.

[16] This form can be used when there is no overt object (in parallel with non-past predicates, see Section 4), i.e., **nân ɗeke dɛ́r Ø.** (1SG CJ.IPFV:want cook 3) 'I want to cook it.'

[17] We do not have any evidence that Bari has a split based on the animacy hierarchy (i.e., pronoun, proper noun, animate/inanimate noun, etc.).

3.1 Topicality of the object

At a first sight, the split seems triggered by definiteness, as indicated by previous studies (Spagnolo 1933; Yokwe 1978). When the object is modified by an anaphoric demonstrative or semantically definite adjectives, the non-suffixed form is preferred, although use of the suffix form is not ungrammatical (15). This implies that definiteness has only indirect correlation with this verbal alternation in Bari.

(15) a. nân a **{mɛ́t / ?mɛd-djâ}** ŋ́ótʊ[18] lɔ́.[19]
1SG PRT **{see / see-SUF}** person PROX/DEF.M
'I saw this/the person.'

b. nân a **{mɛ́t / ?mɛd-djâ}** ŋ́ótʊ lɔ́-gɛ́lɛŋ.
1SG PRT **{see / see-SUF}** person M-only
'I saw the only person'

c. nân a **{mɛ́t / ?mɛd-djâ}** (ŋɔ) lɪŋ.
1SG PRT **{see / see-SUF}** (thing) all
'I saw all/everything.'

Examples (16) and (17) will demonstrate that the split correlates rather directly with information structure. The non-suffixed form is obligatory when the object is the (in-situ) topic of the sentence, that is, the noun phrase about which something is said. The same principle applies when the object is an infinitive verb (coded by the 'passive/infinitive' form), as shown in (18).

(16) <Q1. What did you see in Paris?>

a. nân a **{mɛd-djâ / #mɛ́t}** Eiffel Tower.
1SG PRT **{see-SUF / see}** Eiffel Tower
'I saw the Eiffel Tower.'

<Q2. Did you see the Eiffel Tower in Paris?>

b. je, nân a **{mɛ́t / #mɛd-djâ}** Eiffel Tower.
yes 1SG PRT **{see / see-SUF}** Eiffel Tower
'Yes, I saw the Eiffel Tower.'

[18] Due to a tone sandhi rule, this word will be realized as **ŋ́ótʊ** after **mɛ́t** but **ŋ́ótʊ́** after **mɛd-djâ**. Hereafter, when the tonal form of the object differs according to the alternatives shown in {A/(*)B}, we will represent it in accordance with the candidate A.

[19] The attributive demonstratives appears either before (with deictic meanings) or after the noun (with anaphoric meanings). Contextually, the postnominal proximate demonstratives can function as definite articles.

c. je, nân a **mɛ́t** Ø.
yes 1SG PRT **see** 3
'Yes, I saw it.'

(17) <Q. Where is the chicken I bought from the market?>
a. nân a **{dɛ́r / #dɛ́r-ɟá}** súkuri ná.
1SG PRT **{cook / cook-SUF}** chicken DEF/PROX.F
'I cooked the chicken.'
b. nân a **{dɛ́r / #dɛ́r-ɟa}** Ø.
1SG PRT **{cook / cook-SUF}** 3
'I cooked it.'

(18) <Q1. What did you see?>
a. nân a **mɛd-djâ** dɛ́r-â ná sukúri
1SG PRT **see-SUF** cook-PASS POSS.SG.F chicken(.F)
'I saw the cooking of chicken.'
<Q2. Did you cook the chicken?>
b. áʔa, nân a **mɛ́t** dɛ́r-â ná sukúri
no 1SG PRT **see** cook-PASS POSS.SG.F chicken(.F)
'No, [I did not cook chicken but] I saw [it,] the cooking of chicken.'

As expected, the non-suffixed form is obligatory when the object is left-dislocated (19a), which is semantically nearly identical with the in-situ topic sentence (19b).[20] This distribution is also attested in narrative texts, as contrasted in (20).

(19) a. súkúri ná, nân a **{dɛ́r / *dɛ́r-ɟa}** Ø.
chicken DEF.F 1SG PRT **{cook / cook-SUF}** 3
'The chicken, I ate (it).'
b. nân a **dɛ́r** súkuri ná
1SG PRT **cook** chicken DEF.F
'I ate [it,] the chicken.'

(20) Matat lele kata karen ka-nyit a Nyombe.
chief INDEF.SG.M exist name(.PL) PL-POSS.3SG COP Nyombe
Lepeŋ a **yem-ba** wate ka-nyit gwoso pwök wod buker.
3SG PRT **marry-SUF** wife.PL PL-POSS.3SG like sixteen

[20] According to my consultant's own metalinguistic expression, 'in these sentences, the object sounds like a subject, but it is not.' These two sentences, according to him, signal exactly the same meaning.

Kine wate ka-nyit lepeŋ **yem** i kiŋa-jin ka'de ka'de.
PROX.F.PL wife.PL PL-POSS.3SG 3SG **marry** in year-PL different
'There was a chief whose name was Nyombe. He married about sixteen wives. These wives of his, he married in different years.'
(Lado & Lako 1999: 103, gloss, boldface, hyphen and translation are my own)[21]

In other words, non-topic objects always follow the **-ɟV** form regardless of definiteness, and this is typically the case when the object is a *Wh*-answer (16a) or in the scene-setting context (20). Personal pronouns can co-occur with the **-ɟV** form when it is a *Wh*-answer (21),[22] in contrast with (1c).

(21) <Q. What did you cook, the chicken that I gave you?>
je, nân a **dɛ́r-ɟá** ɲe.
yes 1SG PRT **cook-SUF** 3SG
'Yes, I cooked it.'

However, the **-ɟV** form is not directly correlated with (object) focus (see also Section 4). The object that follows it does not always represent new information (22), and the cleft-like focus construction also requires the non-suffixed form (23).

(22) pɔní a **dɛ́r-ɟá** lókoré nanô?
Poni PRT **cook-SUF** meat when
'When did Poni cook (the) meat?'

(23) súkúri nagwɔn nân a **dɛ́r** Ø.
chicken(.F) F.FOC 1SG PRT **cook** 3
'It is (the) chicken that I cooked.'

Also note that the alternation is not directly related with the information structure of the predicate. If the entire verb phrase including an object is new to the context, the suffixed form is used, but if the verb is new but the object represents the old information, the non-suffixed form is used (24), in parallel with (18) given above.

[21] This example is represented in written Bari register and orthography used in early primary education and slightly different from our consultant's dialect.
[22] This answer can also be **je, nân a dɛ́r Ø.** (yes 1SG PRT cook 3) since the object can be interpreted as the sentence topic (i.e., already introduced by the question). Note that the overt pronoun is obligatory after **-ɟV**.

(24) a. <Q. 'What did the mother do?'>
mamá a **{gwúd-djʌ́ / #gwút}** ŋʊɔ.
mother PRT **{hit-SUF / hit}** child
'The mother hit the child.'
b. <Q. 'Did you cook the chicken?'>
áʔa, nân a **{ɲʌ́ / #ɲé-sú}** súkuri ná.
no 1SG PRT **{eat / eat-SUF}** chicken PROX.SG.F
'No, I *ate* [it,] the chicken.'

From a cross-linguistic viewpoint, it is expected that the suffix form will always be used before an interrogative object (e.g., **ɲɔ̂** 'what,' **ŋâ** 'who'), and this is true for Bari as long as normal conversational questions are concerned (25a). However, Bari allows the non-suffixed form, too (25b), when the sentence is used in a 'review' question, such as in the questions in a children's reader, that is, when the questioner knows the answer (the answer requires the same verbal form). The same applies the question with an indefinite noun **ŋɔ̂** '(some/any)thing' (26). [23] These 'review' questions are unlikely to code contrastivity.

(25) a. pɔní a **dɛ́r-ɟá** ɲɔ?
Poni PRT **cook-SUF** what
'What did Poni cook?' (conversational question)
b. pɔní a **dɛ́r** ɲɔ?
Poni PRT **cook** what
'What did Poni cook?' ('review' question)

(26) a. ɗirí, pɔní a **dɛ́r-ɟá** ŋɔ?
Q Poni PRT **cook-SUF** thing
'Did Poni cook anything?' (conversational question)
b. ɗirí, pɔní a **dɛ́r** ŋɔ?
Q Poni PRT **cook** thing
'Did Poni cook anything?' ('review' question)

[23] This dichotomy seems to be cross-linguistically rare (see Ambar 2008). To supplement the picture, Bari exhibits some more options for *Wh*-questions. Notably, the contrastive question with the 'focus' form of the verb, which contrastively focalizes the post-verbal element (see Section 4): **pɔní dɛr-ɟá ɲɔ?** (Poni FOC:cook-SUF what) 'What [exactly, in contrast to the others] did Poni cook?'. In addition, the interrogative pronouns in the examples in (23) can optionally be fronted with the focus marker (**lɔgwɔn**), i.e., **ɲɔ̂ lɔgwɔn pɔní a dɛ́r(-ɟa)?** (what M.FOC Poni PRT cook-SUF) 'What did Poni cook?' The fronting does not seem to change the meaning.

3.2 Triggers other than the topicality of the object

Although the topicality of the object is apparently the major trigger for the alternation as demonstrated above, the use of the **-ɟV** suffix is obligatory in some other contexts, such as when the verb is followed by non-referential object nouns (27)[24] or a complement clause introduced by **adí** 'that' or **kɔ́** 'if' (28).

(27) a. nân a **mɛd-djâ** tílibiɟôn.
1SG PRT **see-SUF** TV
'I watched TV.' (non-referential)
b. nân a **mɛ́t** tílibiɟôn.
1SG PRT **see** TV
'I saw the TV [device, that you indicated].' (referential, topic)

(28) a. nân a **{wúr-ɟʌ / *wúr}** adí nân pɔ́~pɔ̂.
1SG PRT **{write-SUF / write}** COMP 1SG DJ.IPFV~come
'I wrote that I am coming.'
b. nân a **{pɪ-ɟâ / *pí}** kɔ́ ɲe pɔ́~pɔ̂.
1SG PRT **{ask-SUF / ask}** if 3SG DJ.IPFV~come
'I asked if he is coming.'

Moreover, intentionality, another transitivity-related feature of the predicate, affects it in certain conditions. When the object is a reflexive pronoun, a third person pronoun or a definite noun, the non-suffixed verb can code intentional action (29, 30).

(29) a. nân a **gwút** múgʊn.
1SG PRT **hit** body
'I hit myself (intentionally).'
b. nân a **gwúd-djʌ́** múgʊn.
1SG PRT **hit-SUF** body
'I hit myself (accidentally).'

(30) a. nân a **gwɔ́** {lɔ́pɛŋ / mɔkɔt lɔ́-ɲɪt}.
1SG PRT **kick** {3SG / leg(.M) SG.M-POSS.3SG}
'I kicked him / his leg (intentionally).'
b. nân a **gwɔ́-ɟá** {lɔ́pɛŋ / mɔkɔt lɔ́-ɲɪt}.
1SG PRT **kick-SUF** {3SG / leg(.M) SG.M-POSS.3SG}
'I kicked him / his leg (accidentally).'

[24] As such, example (1b) also means 'I cooked chicken (non-referential)'.

This distinction is only represented in thetic sentences that describe 'out-of-blue' information (e.g., as an answer to the question 'What happened?'). When the object represents the only new information in the sentence, **-ɟV** is required (31).

(31) <Q. Who did you kick?>
nân a **{gwɔ́-ɟá / #gwɔ́}** lɔ́pɛŋ.
1SG PRT **{kick-SUF / kick}** 3SG
'I kicked him (intentionally/accidentally).'

In summary, the non-suffixed form is used either to code topicality or, in marginal contexts, intentionality, and is functionally marked, while the **-ɟV** form is used elsewhere.

3.3 Relativization

Bari has two types of relative markers: The long forms (M. **lɔ́gwɔ̂n**, F. **nágwɔ̂n**) and the short forms (M. **lɔ́**, F. **ná**). The long form is structurally straightforward, as they require a TAM marking (e.g., preterit **a**) and retain the basic S/A-V-O word order. On the other hand, the short form requires the V-S/A-O word order and neutralizes the TAM distinction. Relativized objects require the non-suffixed form (32), while relativized subjects require the suffixed form (33). In this construction, topicality of the object is irrelevant to the alternation. In parallel with the other examples discussed above, the non-suffixed form can be analyzed as having a zero-coded resumptive pronoun.

(32) a. ŋʊ́tʊ́ lɔ́gwɔ̂n a **{dɛ́r-ɟá / *dɛ́r}** kíɲo ná
person(.M) REL.M PRT **{cook-SUF / cook}** food(.F) DEF.F
'a man who cooked the food'
b. ŋʊ́tʊ́ lɔ́ **{dɛ́r-ɟá / *dɛ̂r**[25]**}** kíɲo ná
person(.M) REL.M **{cook-SUF / cook}** food(.F) DEF.F
'a man who cooked/cooks/is cooking/etc. the food'

(33) a. kíɲó nágwɔ̂n pɔní a **{dɛ́r / *dɛ́r-ɟa}** (Ø) na
food(.F) REL.F Poni PRT **{cook / cook-SUF}** (3) DEF.F
'the food that Poni cooked'
b. kíɲó ná **{dɛ̂r / *dɛ́r-ɟá}** pɔní (Ø) na
food(.F) REL.F **{cook / cook-SUF}** Poni (3) DEF.F
'the food that Poni cooked/cooks/is cooking/etc.'

In the adjunct relative clauses, both verbal forms are used (34). The non-suffixed

[25] This form represents the general tone sandhi (**dɛ́r → dɛ̂r**).

form is used when the object is the clause-internal topic. In these examples, the resumptive/anaphoric adverbial **kátá** 'in it' is obligatory.

(34) a. pirît nágwɔ̂n lɔpɛ́ŋ a **{dɛ́r-ɟá / dɛ́r}** kíɲo kátá na
place(.F) REL.F 3SG PRT **{cook-SUF / cook}** food in.it DEF.F
'the place where he cooked (the) food'
b. pirît ná **{dɛ́r-ɟá / dɛ̂r}** lɔ́pɛŋ kíɲó kata ná
place(.F) REL.F **{cook-SUF / cook}** 3SG food in.it DEF.F
'the place where he cooked/cooks/is cooking/etc. (the) food'

The anticausative (35) and antipassive (36) constructions are always relativized with the suffixed form.

(35) a. kɔ́pɔ nágwɔ̂n a **ɓéléŋ-gû**
cup(.F) REL.F PRT **break-SUF**
'a cup that broke'
b. kɔ́pɔ ná **ɓéleŋ-gû**
cup(.F) REL.F **break-SUF**
'a cup that broke/breaks/is breaking/etc.'

(36) a. ŋʊ́tʊ́ lɔ́gwɔ̂n a **dɛ́r-ɟa**
person(.M) REL.M PRT **cook-SUF**
'a man who cooked'
b. ŋʊ́tʊ́ lɔ́ **dɛr-ɟa**
person(.M) REL.M **cook-SUF**
'a man who cooked/cooks/is cooking/etc.'

4. Neutralization of the alternation

The same alternation as the preterit affirmative as discussed above is exhibited with some other preverbal particles such as the perfect **aɟé**, the negative preterit **ákɔ**, the negative perfect **ɲuŋ ákɔ**, as well as an imperative form marked with a suffix (37). The frequentative stem also shows the same pattern (38).

(37) a. **dɛr-ɛ́** Ø.
cook-IMP 3
'Cook it!'
b. **dɛr-ɟ-í.**
cook-SUF-IMP
'Cook!'

c. **dɛr-ɛ́** lókoré.
cook-IMP meat
'Cook the meat.'
d. **dɛr-ɟ-í** lókoré.
cook-SUF-IMP meat
'Cook (some) meat.'

(38) a. nân a **dɛ~dɛ̂r** Ø.
1SG PRT **FREQ~cook** 3
'I used to cook it.' (zero pronoun, topic object)
b. nân a **dɛ~dɛ́r-ɟa.**
1SG PRT **FREQ~cook-SUF**
'I used to cook.' (antipassive)
c. ɲé a **gwu~gwût** ŋórɔ lɔ́.
3SG PRT **FREQ~hit** child DEF/PROX.SG.M
'He hit the/this child.' (topic object)
d. ɲé a **gwu~gwúd-djʌ** ŋâ?
3SG PRT **FREQ~hit-SUF** who
'Who did he hit?' (non-topic object)

However, not all Bari clauses show the alternation. The clearest case is the clauses with a 'focus' verb form (marked by the word-initial low tone; see 2.1). Here, the same suffix is used to distinguish the post-verbal focus form (39, **dɛr-ɟá**) and transitive adjunct focus form (40, **dɛr**).[26] In Bari, focus sentences are TAM-neutral as in the 'short' relative clauses. Both forms code contrastive focus for the postverbal or adjunct element, which is obligatory.

(39) a. nân **dɛr-ɟá** súkuri.
1SG **FOC:cook-SUF** chicken
'I cooked/cook/can cook/etc. (the) chicken [not (the) beans, etc.]' (transitive)
b. nân **dɛr-ɟá** í ɓaŋ.
1SG **FOC:cook-SUF** in house
'I cooked/cook/can cook/etc. in the house [not outside, etc.].' (antipassive)
c. *nân **dɛr-ɟá**.
1SG **FOC:cook-SUF**

[26] There are cleft-like contrastive focus markers (Nakao, forthcoming), which are not discussed here.

(40) a. nân **dɛr** súkúri í ɓaŋ.
1SG **FOC:cook** chicken in house
'I cooked/cook/can cook/etc. (the) chicken in the house [not outside, etc.]'
b. nân **dɛr** Ø í ɓaŋ.
1SG **FOC:cook** 3 in house
'I cooked/cook/can cook/etc. it in the house [not outside, etc.]'
c. *nân **dɛr** súkúri.
1SG **FOC:cook** chicken

The postverbal focus form has developed into an imperfective (habitual/progressive) marker without changing the phonological realization (for transitive dynamic verbs; Nakao, forthcoming). Although it no longer codes contrastive focus, it always requires an object (41), whose topicality does not cause the verbal alternation.[27] We accept the term 'conjoint' for this form from Bantu linguistics, where it signifies the form that cannot occur sentence-finally (Nakao, forthcoming).

(41) <Q1. What are you cooking?>
a. nân **dɛr-ɟá** súkuri.
1SG **CJ.IPFV:cook-SUF** chicken
'I am cooking (the/a) chicken.'
<Q2. Are you cooking the chicken that I brought?>
b. je, nân **dɛr-ɟá** súkuri.
yes 1SG **CJ.IPFV:cook-SUF** chicken
'Yes, I am cooking (the/a) chicken.'
c. je, nân **dɛr-ɟá** ɲe.
yes 1SG **CJ.IPFV:cook-SUF** 3SG
'Yes, I am cooking it.'

In parallel, the non-past form (marked with partial reduplication and optional **-ɟV**) does not show the alternation when followed by an object (42) either.[28]

[27] As such, the object can be left-dislocated: **súkúri ná, nân dɛr-ɟá ɲe.** (chicken.F DEF.F 1SG CJ.IPFV:cook-SUF 3SG) 'The chicken, I am cooking (it).' Spagnolo's observation that the **-ɟV** form is used when 'the action is habitual rather than exclusive' (see Section 1) is based on this construction and 'preterit **a** + **-ɟV**' (e.g., (1b)) does not code past habitual.
[28] The non-past form is interpreted as potential with dynamic verbs.

(42) <Q1. What can you see?>
a. nân **{mɛ~mɛ́t / mɛ́~mɛd-djâ}** Eiffel Tower.
1SG **{NPST~see / NPST~see-SUF}** Eiffel Tower
'I can see Eiffel Tower.'
<Q2. Can you see Eiffel Tower?>
b. je, nân **{mɛ~mɛ́t / mɛ́~mɛd-djâ}** Eiffel Tower.
yes 1SG **{NPST~see / NPST~see-SUF}** Eiffel Tower
'Yes, I can see Eiffel Tower.'
c. je, nân **mɛ~mɛ́t** Ø.
yes 1SG **NPST~see** 3
'Yes, I can see it.'

Negative imperfective and non-past sentences are distinguished by the **-ɟV** suffix again regardless of topicality of the object, where it looks as if a TAM marker (43).

(43) a. nân ti **dɛ́r-ɟá** súkuri.
1SG NEG **cook-SUF** chicken
'I am not cooking (the) chicken.' (imperfective)
b. nân ti **dɛ́r** súkuri.
1SG NEG **cook** chicken
'I cannot cook (the) chicken.' (potential)

The imperfective/potential dichotomy is neutralized when the verb occurs sentence-finally, which, to use the Bantu linguistics term, involves the 'disjoint' form (44). At a first sight, (42c) and (44b) may seem to represent the same alternation demonstrated by (1c,d) and to perhaps be diachronically related, but the synchronic picture is, as such, more complex.[29]

(44) a. nân **dɛ́~dɛr-ɟa**.
1SG **NPST/DJ.IPFV~cook-SUF**
'I can cook/am cooking.'
b. nân **mɛ́~mɛd-djâ**.
1SG **NPST/DJ.IPFV~see-SUF**
'I can see (I am visible)/I am seeing.'

[29] Bari has some other TAM markers such as **dê** (future), **kɔ́/tô** (perhaps posterior), **kʌɟû** (past) placed before IPFV/NPST verb forms, which basically parallel the patterns described here.

In addition, there are several marginal TAM markings, such as in purpose clauses (45a), sequential clauses (45b) and inchoative aspect (45c), where the **-ɟV** suffix plus the 'imperative/sequential' suffix (with falling tone) is obligatory and the verbal alternation is not attested (***wur-ê**, ***dɛr-ɛ̂** and ***dɛr-ɛ̂ dɛ́r**, respectively).

(45) a. lɔpɛ́ŋ dɛ̀kɛ gálam aɲɛ́n ɲe **wur-ɟ-î** karín ka-ɲɪt.
3SG CJ.IPFV:want pen PURP 3SG **write-SUF-SEQ** name(.PL) PL-POSS.3SG
'He wants a pen so that he can write his name.'

b. ɲé a tô ɪ ɓâŋ, á ɲe **dɛr-ɟ-î** súkúri ná.
3SG PRT go to house then 3SG **cook-SUF-SEQ** chicken DEF.F
'He went home and then cooked the chicken.'

c. lɔpɛ́ŋ **dɛr-ɟ-î** **dɛ́r-ɟá** kíɲo.
3SG **cook-SUF-SEQ** **cook-SUF** food
'Now he cooks food [as a new habit].'

5. Diachrony and contact

The previous sections demonstrated that **-ɟV** is a multifunctional, or rather a functionally unmarked suffix for transitive verbal roots. This markedness paradox will be approached from comparative-historical and language contact perspectives in this section.

Comparative evidence suggests that the Bari voice-marking suffixes were derived from (later) proto-southern Eastern Sudanic with possible cognate forms attested in the Nilotic, Surmic and Daju branches.[30] Maasai seems to have the richest voice-marking morphology within Eastern Nilotic, where a transitive verb may have antipassive, anticausative and passive markings (46).

(46) Maasai (Tucker & Mpaayei 1955: 135, 139; transcription and labels are my own)

a.	active	a-rapoʃ	'I fill it.'[31]
	antipassive	a-rapoʃ-**iʃo**	'I fill.'
	anticausative	a-rapoʃ-**o**	'I am filled.'
b.	anticausative	ɛ-gɪl-**a**	'it gets broken by itself, it breaks.'
	passive	ɛ-gɪl-**ɪ**	'it is broken by somebody'

[30] See Dimmendaal and Jakobi (2020) for the north-south split within Eastern Sudanic. Northern Eastern Sudanic, which has AOV word order, attests to the orthodox DOM (Dimmendaal 2010), while it is probably irrelevant to DOM in Bari. Some other Nilo-Saharan languages have antipassive (Tucker & Bryan 1966), but they seem to be etymologically unrelated to the Nilotic-Surmic-Daju antipassive suffix. For now, we do not have enough data for the Eastern Jebel and Temein branches within the southern Eastern Sudanic branch.

[31] Like Bari, Maasai has a zero pronoun for the third person.

If we heuristically assume that Maasai retains the archaic system, we can suppose a chain-like shift in Bari; that is, old passive (***-I**) and anticausative (***-A/-O**) merged as passive, while antipassive (Maasai **-iʃo/-ɪʃɔ**) came to cover anticausative and other functions (**-ɟV**). Note that Bari has an irregular verb 'eat,' whose suffixes correspond to the regular Maasai suffixes (Maasai data are adapted from Mol 1996).[32]

(47)

Maasai	Bari	
-ɲa	ɲʌ́	'eat' (active)
-ɲa-**ɪʃɔ**	ɲé-**su**	'do eating' (antipassive)
-ɲa-**ɪ**	ɲʌ́-**jî**	'be eaten' (passive)

Southern Nilotic Nandi also has a phonologically similar antipassive suffix **-i**:**s/-sa**, while it lacks passive and anticausative constructions (Creider & Creider 1989: 92). Apart from Nilotic, Tennet (Surmic) has antipassive **-ye** versus passive **-e/-o** (see also Schröder 2006).

(48) Tennet (adapted from Randal 1998: 245)

a. á-**dáh** doléc áhát
IPFV-**eat** child.NOM asida
'The child is eating asida [polenta].'

b. á-**dáh-ye** doléc
IPFV-**eat-AP** child.NOM
'The child is eating.'

c. á-**dáh-é** ahat-a balwáz
PERF-**eat-PASS** asida-NOM yesterday
'The asida was eaten yesterday.'

Moreover, Shatt (Caning), a Daju language spoken in the Nuba Mountains in Sudan, attests reflexive **-wa**, antipassive **-za** and an antipassive-like 'continuous' suffix **-VccV** that deletes the object when attached to a transitive verb (Alfira *et al*. 2016). Tucker & Bryan (1966: 239) give antipassive suffixes **-VcV** (imperfective) and **-si** (perfective) in Shatt and hint at their link with Eastern/Southern Nilotic antipassive.

While the reconstruction of an antipassive suffix with a certain palatal consonant seems justifiable for the common ancestor of Nilotic, Surmic and Daju branches, the functional change that occurred in Bari needs more explanation. It is noteworthy that

[32] The correspondence of Bari **-ɟV**, Maasai **-iʃ(o)** and Southern Nilotic Nandi **-isye/-se** was first noticed by Murray (1920), who also considered the Nubian suffix **-j/-ch** ('intensive and reflexive') cognate (Murray 1923). Payne (2021) compares the Maasai suffix with the verb **ɪʃɔ(r)** 'give,' while she also notes its phonological similarity with Bari and Southern Nilotic.

some Nilotic languages attest parallel syncretism of the old voice-marking affixes.

Eastern Nilotic Lopit has a multifunctional **-a/-o** for anticausative, antipassive and imperfective (Moodie & Billington 2020, who consider these to be different suffixes), although its antipassive usage does not allow direct object. Not only that the functional range that this suffix covers is similar to Bari **-ɟV**, Bari has irregular transitive roots such as **ɗúk** 'build' (49), which has **ɗúk-â** as a free variation of **ɗúg-gʌ** (homophonous with the passive **ɗúk-â**).

(49) a. nân a {**ɗúk-â** / **ɗúg-gʌ́**} kadî.
1SG PRT {**build-SUF** / **build-SUF**} house
'I built a house.' (active)

b. kadî a {**ɗúk-â** / ***ɗúg-gʌ**}.
house PRT {**build-PASS/SUF** / **build-SUF**}
'The house is/was built.' (anticausative, passive)

c. nân {**ɗuk-ʌ́** / **ɗug-gʌ́**} kadî.
1SG {**CJ.IPFV:build-SUF** / **CJ.IPFV:build-SUF**} house
'I am building a house.' (imperfective)

In parallel to Lopit, the Northern Lwo languages of the Western Nilotic branch, including Luwo, Shilluk, Anywa and Päri, share a functionally vague suffix **-o/-ɔ**, which appears in antipassive, imperfective and infinitive forms (Andersen 1988; Reh 1996; Storch 2014; Remijsen & Ayoker 2018).[33] In Luwo, this suffix also appears in the anticausative stem (Storch 2014).

(50) Luwo (Storch 2014: 196–197)

a. kádò à-**tháàl**=ˈé
broth:SG PFV-**cook**=3SG
's/he cooked broth'

b. à-**tɛɛdɔ̀** ké kádò
PFV-**cook:AP** PREP broth:SG
's/he cooked some broth'

(51) Luwo (Storch 2014: 125)

a. mɔ́ɔ́gè à-ké-**ɲʌ́mɔ̀**
some:PL PFV-DUR-**eat:AP**
'some were eating'

[33] The morphological alternation also involves a [±ATR] vowel alternation. Päri and Anywa have a passive suffix **-i**, which resembles Maasai (Andersen 1988; Reh 1996).

b. mɔ́ɔ́gè à-ké-**ɲámò**
some:PL PFV-DUR-**eat:AC:AP**
'some were eaten'

Southern Lwo languages, which are more innovative than Northern Lwo and are spoken in the vicinity of the Bari-speaking areas, deserve special attention since the **-o/-ɔ** suffix has become the default (typically transitive, but not exclusively so) verbal ending in this group.[34] Kohnen (1933: 122) implies that the Northern Lwo antipassive form has become the default verbal stem in Southern Lwo. This explanation is supported by the fact that the Northern Lwo antipassive forms, such as 'cook', have become the common transitive stem in Southern Lwo (50, 52). Since Northern Lwo languages are (or at least resemble) ergative type languages, we can consider that this antipassive-to-active shift accompanied ergative-to-accusative shift, as is cross-linguistically recurrent (see Dixon 1994; Harris & Campbell 1995; Aldridge 2011; Hemmings 2021).

(52) Lango (Southern Lwo, Noonan 1992: 132)
lócà ò-**tèdò** rìŋó
man 3SG-**cook.PERF** meat
'the man cooked the meat'

Although no known Central Sudanic language neighboring Bari attests to an antipassive construction, Lulubo, an exceptional language of the Moru-Ma'di group spoken in an area adjacent to Bari, deserves attention. This language exhibits a unique imperfective transitive construction, which (for vowel-initial verbs) requires the post-verbal 'transitive' marker **rʊ̄** ~ **rī** (Andersen 1994). This transitive marker is homophonous with the reflexive pronoun and as such resembles Bari and Southern Lwo in that a de-transitive marker have developed into a transitive/active.

Finally, we can add to this areal replication tendency the restructuring of Juba Arabic (an Arabic creole; see Nakao 2012, 2015, 2017) verbal system, which is characterized by the verbal ending **-u** (typically transitive, but not exclusively so, like Southern Lwo **-o**) and the prevalence of zero-coded causative-anticausative alternation (53).

(53) Juba Arabic
a. ána **rákabu** gidéda.
1SG **cook** chicken
'I cooked (the/a) chicken.' (active)

[34] Although the **-o** suffix is generalized as the verbal ending in Southern Lwo, some members (notably Kenyan Luo) also exhibit a [±ATR] alternation for antipassive.

b. ána **rákabu**.
1SG **cook**
'I cooked (did cooking).' (antipassive)
c. gidéda **rákabu**.
chicken **cook**
'The/a chicken is cooked (state).' (anticausative)
d. gidéda **rakabú**.
chicken **cook:PASS**
'The/a chicken was cooked (by someone).' (passive)

The lack of voice-alternation in Juba Arabic is symbolic because the other traditional Arabic varieties share it. As such, this pattern must be explained in terms of language contact and creolization (with Bari as an important substrate). Another interesting point is that Juba Arabic borrows Bari verbs in the **-ɟV** form (e.g., Juba Arabic **kúr(u)ɟu** 'cultivate', **túr(u)ɟu** 'chase' < Bari **kɔr-ɟô**, **tur-ɟâ**), which would imply that the speakers consider the suffixed form to be the default.

6. Conclusion

This study described the morphosyntactic alternation in Bari where the marked verb form (with **-ɟV**) has a far wider distribution than the unmarked form. The latter typically codes object topicalization and as such has a marked function. Given this paradoxical functional and morphological distribution, this alternation can be portrayed as a typological pivot between the two morphosyntactic phenomena, AP and DOM, which have until recently not been deemed to be so closely related in the literature.

We explained this markedness paradox from a diachronic perspective, that is, as a functional shift of the **-ɟV** suffix from antipassive to general active/transitive, representing a historical pivot of AP and DOM. This type of shift is attested, especially along the process of 'ergative-to-accusative' alignment change, which Southern Lwo languages may have followed, but the Bari case is unique since it occurred in a (syntactically) accusative language. We have also discussed the possibility that the development of Bari **-ɟV** as an areal phenomenon comparable with the morphological shift in some neighboring languages. Since the morphological development of Nilo-Saharan languages have not been investigated in depth so far, this study contributes to the historical linguistic understanding of this largely understudied language family.

Abbreviations

1, 2, 3	person	DEF	definite article
AC	anticausative	DJ	disjoint
AP/AP	antipassive	DOM	differential object marking
CJ	conjoint	DUR	durative
CL.7	class 7	F	feminine
COMP	complementizer	FOC	focus
COP	copula	FREQ	frequentative
IMP	imperative	NOM	nominative
INDEF	indefinite	NPST	non-past
IPFV	imperfective	OBJ	object
M	masculine	PASS	passive
NEG	negative	PERF	perfect
PFV	perfective	PURP	purpose clause marker
PL	plural	REL	relative clause marker
POSS	possessive	SEQ	sequential
PREP	preposition	SG	singular
PROX	proximate	SUF	suffix
PRT	preterit	TAM	tense-aspect-modality

Acknowledgement

This study is funded by JSPS KAKENHI (Grant Number 19K13160, 19H01254, 23K21931, 24K16054). The author appreciates Dr. David Gore for providing Bari and Juba Arabic data and for his friendship. An early version of this study was presented at the 59th Japanese Association of African Studies conference on May 21, 2022 (online, Nagasaki University). The author thanks all who commented on it at the conference.

References

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21: 435–483.

Aldridge, Edith. 2011. Antipassive in Austronesian alignment change. In: Dianne Jonas, John Whitman & Andrew Garrett (eds.) *Grammatical change: Origins, nature, and outcomes*. Oxford: Oxford University Press. pp. 332–346.

Alfira, David Abbi, Timothy Kuku Kafi, Hassan Kuwa Kaki, Ali Alaliim Hasan, Anjo Kuku Anjo, Dayan Kuku Jas & Sadik Kafi Sarukh. 2016. *Caning grammar book* (trial edition). Juba: Sudan Workshop Programme.

Ambar, Manuela. 2008. Wh-asymmetries. In: Anna Maria Di Sciullo (ed.) *Asymmetry in grammar. Volume I: Syntax and semantics*. Amsterdam: John Benjamins. pp. 209–

249.
Andersen, Torben. 1988. Ergativity in Päri, a Nilotic OVS language. *Lingua* 75: 289–324.
Andersen, Torben. 2018. The encoding of subjects and objects in Jumjum, a Nilotic OV language. *Lingua* 204: 78–116.
Ashton, E. O. 1944. *Swahili grammar (including intonation)*. London: Longmans, Green & Co.
Bárány, András. 2012. Hungarian conjugations and differential object marking. In: Balázs Surányi & Diána Varga (eds.) *Proceedings of the first Central-European conference for postgraduate students*. Piliscsaba: Pázmány Péter Catholic University. pp. 3– 25.
Cohen, Kevin Bretonnel. 2000. *Aspects of the grammar of Kukú*. Munich: Lincom.
Creider, Chet A. & Jane Tapsubei Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske.
Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure.* Cambridge: Cambridge University Press.
Dimmendaal, Gerrit J. 2010. Differential object marking in Nilo-Saharan. *Journal of African Languages and Linguistics* 31: 13–46.
Dimmendaal, Gerrit J. & Angelika Jakobi. 2020. Eastern Sudanic. In: Vossen, Rainer & Gerrit J. Dimmendaal (eds.) *The Oxford handbook of African languages*. Oxford: Oxford University Press. pp. 392–407.
Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
Harris, Alice & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
Hemmings, Charlotte. 2021. When an antipassive isn't an antipassive anymore: The actor voice construction in Kelabit. In: Katarzyna Janic & Alena Witzlack-Makarevich (eds.) *Antipassive: Typology, diachrony, and related constructions*. Amsterdam: John Benjamins. pp. 579–620.
Hieda, Osamu. 2009. Kumamu-go no chūdōsō (in Japanese: Middle in Kumam). *Asian and African languages and linguistics* 4: 39–59.
Janic, Katarzyna & Alena Witzlack-Makarevich. 2021. The multifaced nature of the antipassive construction. In: Katarzyna Janic & Alena Witzlack-Makarevich (eds.) *Antipassive: Typology, diachrony, and related constructions*. Amsterdam: John Benjamins. pp. 1–40.
Kohnen, Bernardo. 1933. *Shilluk grammar, with a little English-Shilluk dictionary*. Verona: Missioni Africane.
König, Christa. 2008. *Case in Africa*. Oxford: Oxford University Press.
Lado, Patrick & Sejario Lako. 1999. *Jujumbu kendya ko Bari* (Buk tomusala, buk ludukötyo). 9th edition, revised. Nairobi: Institute of Regional Languages.
Malchukov, Andrej L. 2017. Ergativity and differential case marking. In: Jessica Coon, Diane Massam & Lisa deMena Travis (eds.) *The Oxford handbook of ergativity*.

Oxford: Oxford University Press. pp. 253–278.
Mol, Frans. 1996. *Maasai language & culture: dictionary*. Lemak: Maasai Centre Lemak.
Moodie, Jonathan & Rosey Billington. 2020. *A grammar of Lopit: An Eastern Nilotic language of South Sudan*. Leiden: Brill.
Murray, G. W. 1920. The Nilotic languages: A comparative essay. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland* 50: 327–368.
Murray, G. W. 1923. *An English-Nubian Comparative Dictionary*. London: Oxford University Press.
Nakao, Shuichiro. 2012. Revising the substratal/adstratal influence on Arabic creoles. In: Osamu Hieda (ed.) *Challenges in Nilotic linguistics and more: Phonology, morphology and syntax*. Fuchu: ILCAA. pp. 127–149.
Nakao, Shuichiro. 2015. Causative-inchoative alternation in Arabic creoles: An 'exceptional' language change? In: Osamu Hieda (ed.) *Information Structure and Nilotic Languages*. Fuchu: ILCAA. pp. 225–239.
Nakao, Shuichiro. 2017. A grammar of Juba Arabic. Doctoral dissertation, Kyoto University.
Nakao, Shuichiro. forthcoming. The origins of the conjoint/disjoint alternation in Bari. manuscript.
Noonan, Michael. 1992. *A grammar of Lango*. Berlin: Mouton de Gruyter.
Payne, Doris L. 2021. The profile and development of the Maa (Eastern Nilotic) antipassive. Katarzyna Janic & Alena Witzlack-Makarevich (eds.) *Antipassive: Typology, diachrony, and related constructions*. Amsterdam: John Benjamins. pp. 447–480.
Polinsky, Maria. 2017. Antipassive. In: Jessica Coon, Diane Massam & Lisa deMena Travis (eds.) *The Oxford handbook of ergativity*. Oxford: Oxford University Press. pp. 308–331.
Randal, Scott. 1998. A grammatical sketch of Tennet. In: Gerrit J. Dimmendaal (ed.) *Surmic languages and cultures*. Cologne: Rüdiger Köppe. pp. 219–272.
Reh, Mechthild. 1996. *Anywa language: Description and internal reconstructions*. Cologne: Rüdiger Köppe.
Remijsen, Bert & Otto Gwado Ayoker. 2018. *A descriptive grammar of Shilluk*. Honolulu: University of Hawaii Press.
Sansò, Andrea. 2017. Where do antipassive constructions come from? A study in diachronic typology. *Diachronica* 34(2): 175–218.
Schikowski, Robert & Giorgio Iemmolo. 2015. Commonalities and differences between differential object marking and indexing (Zurich: University of Zurich). Zenodo. https://doi.org/10.5281/zenodo.7738049
Schröder, Helga. 2006. Antipassive and ergativity in Western Nilotic and Surmic. *Annual Publications in African Linguistics*. 4: 91–110.

Schröder, Helga. 2015. Alignment systems and passive-antipassive distribution in Nilotic languages. *The University of Nairobi Journal of Language and Linguistics* 4: 42–81.
Sinnemäki, Kaius. 2014. A typological perspective on Differential Object Marking. *Linguistics* 52(2): 281–313.
Spagnolo, Lorenzo M. 1933. *Bari grammar*. Verona: Missioni Africane.
Storch, Anne. 2014. *A Grammar of Luwo: An anthropological approach*. Amsterdam: Amsterdam: John Benjamins.
Tucker, Archibald N. & Margaret A. Bryan. 1966. *Linguistic analyses: The Non-Bantu languages of North-Eastern Africa*. London: Oxford University Press.
Tucker, Archibald N. & J. Tompo Ole Mpaayei. 1955. *A Maasai grammar*. London: Longmans, Green & Co.
Vossen, Rainer. 1982. *The Eastern Nilotes: Linguistic and historical reconstructions*. Berlin: Dietrich Reimer.
Witzlack-Makarevich, Alena & Ilja A. Seržant. 2018. Differential argument marking: Patterns of variation. In: Ilja A. Seržant & Alena Witzlack-Makarevich (eds.) *Diachrony of differential argument marking*. Berlin: Language Science Press. pp. 1–34.
Yokwe, Eluzai Moga. 1978. Bari phonology. MA thesis, University of Khartoum.
Yokwe, Eluzai Moga. 1987. The tonal grammar of Bari. Doctoral dissertation, University of Urbana-Champaign.

Descriptive notes on negation particles in Kilimanjaro Bantu languages[1]

Daisuke Shinagawa
Research Institute for Languages and Cultures of Asia and Africa,
Tokyo University of Foreign Studies

Abstract

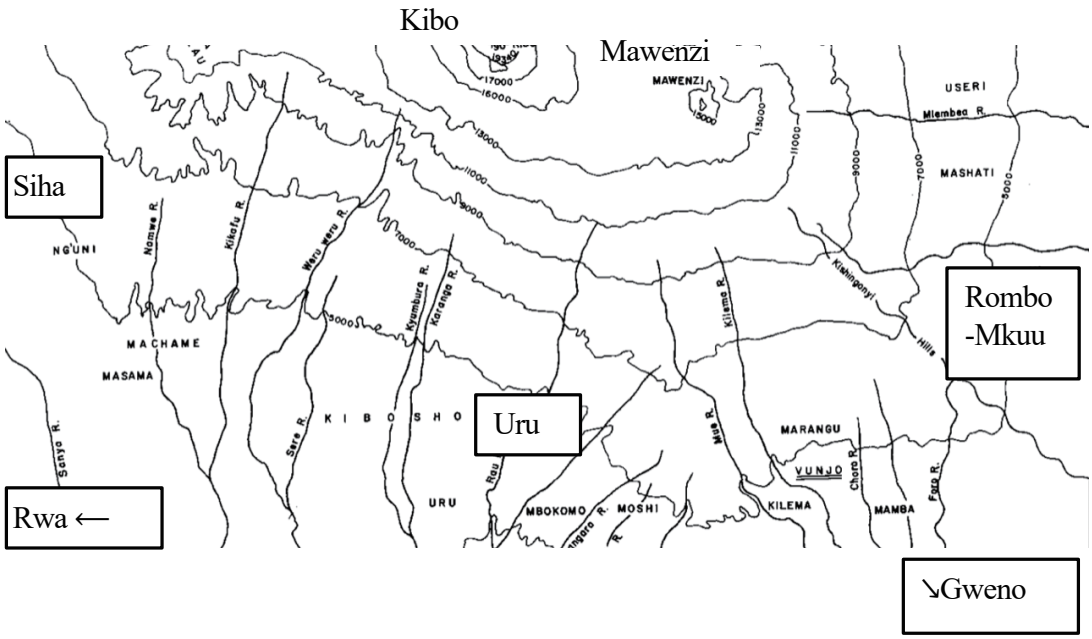
This paper provides descriptive information about the morphosyntactic variation of main-clause negation marking systems in Kilimanjaro Bantu (KB) languages. Findings presented in these descriptive notes include the hybrid marking system found in Uru [E622D], where main-clause negation can be marked either by the invariant marker **pfo**, which is assumed to originate from class 17 locative demonstrative and shared widely in Central Kilimanjaro Bantu languages, or by variable particles which are coreferential with inflectional properties of the subject, which, however, can only be used in the case of non-speech participant subjects. The existence of this system, which can be located on the scale of systematic complexity between the full agreement system attested in Gweno and the invariant marking system attested, e.g., in many varieties of Rombo, provides several insights into the diachronic process of the diversification of main-clause negation marking systems in KB, which are typologically unique and diverse from a cross-Bantu typological perspective.

Keywords: negation, particle, Chaga, Kilimanjaro Bantu

[1] I would like to express my gratitude to all the language consultants who have patiently shared with me their knowledge about their native languages, namely Mzee Afitwa Abia Ayo for Rwa, Mzee Noah Mmari for Siha, Mzee Nikolausi Tailo for Rombo, and Mzee Kalisti Kilewo for Uru. My sincere thanks also go to Benji Wald for sharing his data on Rombo negatives based on the fieldwork he carried out in the 1970s. I hereby acknowledge the following projects that have supported this work in one way or another: JSPS bilateral project "The past and present of Bantu languages: Integrating micro-typology, historical-comparative linguistics and lexicography" (PI: Daisuke Shinagawa and Koen Bostoen, #120212302), the ILCAA joint research project "Diachronic Perspectives on Language Description and Typology in Bantu" (PI: Makoto Furumoto), JSPS grant-in-aid #22720158, and #25770149 (for Young Scientists (B)), #16K02630, #19K00568, and #23K25319 (Grant-in-Aid for Scientific Research (C)). The usual disclaimers apply.

1. Introduction

This article is meant to provide descriptive information about the morphosyntactic variation of negation marking systems attested in Kilimanjaro Bantu (KB) languages. KB is spoken in the outskirt areas surrounding Mt. Kilimanjaro, and most of the varieties are classified into Guthrie's (1967–1971) Chaga group [E60].



Map: Geographical distribution of the KB languages of the present study
(based on Nurse 1979: 58)

Philippson and Montlahuc (2003) identify the following varieties as KB languages, which are classified into three major subgroups, namely Western Kilimanjaro (WK), Central Kilimanjaro (CK), and Rombo, as shown in Table 1. The referential codes given in the table are based on Maho's (2009) update of Guthrie's (1967–71) original codes, which are adopted in the inventory of the Bantu languages compiled by Hammarström (2019).

Table 1: Classification of the Kilimanjaro Bantu languages (cf. Philippson & Montlahuc 2003; Maho 2009; Hammarström 2019)

E621 WK	A	Rwa/Meru	E622 CK	A	Mochi	E623 Rombo	A	Useri
	B	Machame		B	Mbokomu		B	Mashati
	C	Siha		C	Vunjo		C	Mkuu
	D	Kiwoso		D	Uru		D	Keni
	E	Masama						
	F	Ng'uni						
E63		Rusha /Okuma						
E64		Kahe						
E65		Gweno						
E74	a	Dabida						

As is well recognised in the literature (cf. Nurse 1979; Philippson & Montlahuc 2003), main-clause negation marking in KB is generally characterised by the use of clause-final particles and the lack of the negation marking prefix that appears in the pre-initial (PREIN) slot in the common template of the morphological structure of Bantu verbs shown in Table 2.

Table 2[2]: Morphological template of Bantu verbs (cf. Meeussen 1967; Güldemann 2022)

	Pre-stem markers				Stem cluster				
Slot labels	PREIN	IN	POSTIN	PRERAD	RAD	EXT	PREFIN	FIN	POSTFIN
Index no.	-4	-3	-2	-1	0	1	2	3	4
Typical functions	NEG, TAM, etc.	SCd	NEG, TAM, etc.	OCd	Root	Voice	TAM	TAM	Clause type, etc.

Since in Eastern Bantu languages main-clause negation is generally achieved through pre-initial markers, typically those traced back to the Proto-Bantu (PB) reconstructed form ***ka-**, its absence, along with the use of clause-final particles instead, is one of the relatively uncommon structural features from a cross-Bantu viewpoint (cf. Kamba-Muzenga [1978], but see also Devos & van der Auwera [2013] for numerous attestations of verb-external negative particles especially in Western Bantu languages and their

[2] The abbreviations in the table stand for the following: PREIN = pre-initial, IN = initial, POSTIN = post-initial, PRERAD = pre-radical, RAD = radical, EXT = extension (derivational suffix), PREFIN = pre-final, FIN = final (inflectional suffix), POSTFIN = post-final, NEG = negation, TAM = tense, aspect, and modality, SCd = subject concord, OCd = object concord.

typologically common grammaticalisation processes).[3]

From a KB-internal point of view, Nurse (1979, 1981) provides an overview of the negation systems across KB, while Philippson and Nurse (2000) describe a typologically interesting system where negation markers regularly agree with inflectional properties of the subject in Gweno. However, much of the variation observed in KB is still left to be documented in the literature. On the other hand, Devos and van der Auwera (2013) include a dedicated section on the structural variation and its historical development process across KB, to which this paper aims to provide supplementary data from other KB varieties that are not covered by their study, which may in turn lead to a more solid understanding about the historical process by filling in some of the missing pieces of information for their hypothesis.

This article thus aims to provide firsthand descriptive information[4] on the negation marking systems from a wide range of KB varieties, which is intended to be part of the empirical basis for further investigation into the historical process of diversification within KB languages.

2. Variation of main-clause negation marking systems in KB

As shown in the following, considerable variation of negation marking patterns are attested across KB languages. Note, however, that our concern here focuses on main-clause negation; that is, negation in non-main-clauses, or in syntactically dependent clauses, including relative, subjunctive, and infinitive verb forms, is outside the direct scope of this description. As in many Bantu languages, negation in such dependent clauses is consistently marked by verb-internal markers in the pre-initial position (the -2 slot in Table 2) throughout KB varieties, and thus the attested morphosyntactic patterns are cross-linguistically less diversifying.

2.1 Gweno: A full agreement system

As briefly mentioned above, Gweno shows a cross-linguistically complicated system of main-clause negation marking. According to Philippson and Nurse (2000), negative particles in Gweno are extensively coreferential with the subject except for the case of the 2SG subject, as illustrated in (1).

[3] Shinagawa and Marten (2021) show that 35% of their sample languages (20 languages out of 57) utilise at least one negative particle. This percentage is largely comparable to that provided by Nurse (2008: 182–183, 289), who uses a sample of 100 languages, out of which 33 languages have an obligatory negative particle. It's not extremely rare, but still the distribution of the verb-external negation marking is restricted in Eastern Bantu languages. Also see footnote 12 for more information about the distribution of those derived from locative cl. 17 demonstratives.

[4] All the examples provided in this paper, otherwise noted, from my fieldwork.

(1) Main-clause negation marked by full agreement clause-final particles in Gweno (cf. Philippson & Nurse 2000; Devos & van der Auwera 2013)[5]

a. θíleβóníre mndu **ɲí**
θi-le-βon-ire m-ndu **ɲi**
SM$_{1SG}$.NEG-PST.R-see-PST.R 1-entity **NEG$_{1SG}$**
'I didn't see anybody' (Philippson & Nurse 2000: 257)

b. kúʧié **pfó**
ku-Ø-ʧ-ie **pfo**
SM$_{2SG}$-PST.N-come-PST.N **NEG$_{17}$**
'You didn't come' (Devos & van der Auwera 2013: 253)

c. fwáʃiɣa ipfwá **fwé**
fu-a-ʃiɣ-a i-pfwa **fwe**
SM$_{1PL}$-PRS-want-FV INF-die **NEG$_{1PL}$**
'We don't want to die' (Philippson & Nurse 2000: 257)

d. ḿʧiréra mwé
m-ʧi-rer-a **mwe**
SM$_{2PL}$-HAB-talk-FV **NEG$_{2PL}$**
'You don't usually talk' (Devos & van der Auwera 2013: 253)

e. mká θo kwáku áleɣíle mwána **wé**...
m-ka θo kwaku a-le-ɣile mw-ana **we**
1-woman DEM$_1$ because SM$_1$-PST.R-have.PRF 1-child **NEG$_1$**
'that woman, because she had no child' (Philippson & Nurse 2000: 257)

f. βáleβóníre βaɣaŋga **βó**
βa-le-βon-ire βa-ɣaŋga **βo**
SM$_2$-PST.R-see-PST.R 2-witchdoctor **NEG$_2$**
'They didn't see witch-doctors' (Philippson & Nurse 2000: 257)

g. kyumbá ki kínakulwa **kyó**
ki-umba ki ki-na-kul-w-a **kyo**
7-room DEM.P.7 SM$_7$-NOT.YET-sweep-PASS-FV **NEG$_7$**
'This room hasn't been swept yet' (Philippson & Nurse 2000: 263)

Philippson and Nurse (2000) point out that for all classes except for cl. 1 and speech participant subjects, the negative particles are identical in segmental shape to the 'demonstrative of reference', which is morphologically analysed into the stem **-o** preceded by a pronominal prefix, as illustrated in (1f) **βo**[6] {ba-o} and in (1g) **kyo** {ki-o}.

[5] Glosses are given by the present author based on the information provided in Philipson and Nurse (2000).

[6] According to Devos and van der Auwera (2013: 253), and as in usual cases in KB in general, the same form is used as an independent personal pronoun for class 2 nouns as well.

For the speech participant subjects and cl. 1, negative particles are segmentally identical with corresponding personal pronouns except for 2SG, thus 1SG **ɲí**, 1PL **fwé**, 2PL **mwé**, and cl. 1 **wé**, as in (1a), (1c), (1d), and (1e), respectively.

What is arguable here is how to identify 2SG **pfó** in the synchronic system and its origin, which is discussed in Devos and van der Auwera (2013: 252–254). Though it is not a central concern of this paper, at least based on the systematic correspondences across KB languages, it seems to be feasible to propose that the negative particle **pfo** seems to originate from (an extended use of) the locative cl. 17 form, as will be discussed in the following sections.

Finally, it should be mentioned that the contrast between positive and negative verb forms can also be distinguished through tonal alternation applied to the negation particle, as illustrated in (2), quoted from Philippson and Nurse (2000: 257).

(2) a. βáleβóníre βaɣaŋgá **βo**
βa-le-βon-ire βa-ɣaŋga **βo**
SM$_2$-PST.R-see-PST.R 2-witch_doctor **DEM2**
'they saw those witch-doctors'

b. βáleβóníre βaɣaŋga **βó**
βa-le-βon-ire βa-ɣaŋga **βo**
SM$_2$-PST.R-see-PST.R 2-witch_doctor **NEG**
'they didn't see witch-doctors'

Note, however, that the tonal contrast seems not to be extensive in Gweno. According to Philippson and Nurse (2000: 13), tonal contrast in the subject marker slot is only available in the case of speech participant subjects which are realised as low in affirmative and as high in negative clauses.[7] This is substantially different from other varieties where the tonal contrast between negative and affirmative is extensively marked (typically in the verb initial position) as shown in the following sections.

2.2 Mkuu (Rombo): Marking through the invariable particle ku

The other extreme, in terms of the complexity of the negation marking system, is found in Rombo varieties,[8] where main-clause negation is consistently marked by the

[7] Philippson and Nurse (2000: 13) "Basically, in positive non-dependent tenses, [1SG], [1PL], [2SG] and [2PL] prefixes are L, all the others H. In relative and other dependent forms, all subject prefixes are L; in negative forms all are H." (N.B. labels in [ ] are modified by the present author)

[8] However, the situation seems not to have been as simple as illustrated in the examples in (3), which are elicited data only covering canonical main clauses. According to the field notes by Benji Wald, who had a recording session in the 70s with a Rombo speaker who was in his 30s at that time, it seems that not only **ku** but also **fo**, which is an apparent cognate with **pfo** in many CK

invariant marker **ku**.

(3) Main-clause negation marked by the invariable particle **ku** in Rombo-Mkuu (Shinagawa 2015)

a. ŋgiʃeṃlolja ɾafiki jákwa ŋgámá **ku**
ŋgi-ʃe-ṃ-loli-a Ø-ɾafiki i-akwa N-gama **ku**
SM1SG-FUT.R-OM1-see-FV 9-friend 9-POSS1SG 9-tomorrow **NEG**
'I will not see my friend tomorrow'

b. usoma kitabú **ku**
u-Ø-som-a ki-tabu **ku**
SM2SG-PRS-read-FV 7-book **NEG**
'You do not read a book'

c. dukufúnsa kiɾómbó **ku**
du-Ø-kufuns-a ki-ɾombo **ku**
SM1PL-PRS-learn-FV 7-Rombo **NEG**
'We do not learn the Rombo language'

d. múʃeɾunda kasí luní **ku**
mu-ʃe-ɾund-a Ø-kasi Ø-luni **ku**
SM2PL-FUT.R-do-FV 9-work 9-today **NEG**
'You (pl.) will not work today'

e. esakwa héwá **ku**
e-Ø-sakw-a Ø-hewa **ku**
SM1-PRS-watch-FV 9-sky **NEG**
'S/he does not watch the sky'

f. βeʃelóliá **ku**
βe-ʃe-loli-a **ku**
SM2-FUT.R-see-FV **NEG**
'They will not see'

As shown in all the examples in (3), a high tone is consistently assigned to the TBU that immediately precedes **ku**, and spreads even further leftward as illustrated in (3a), (3c), (3e), and (3f), which suggest that the high tone is underlyingly assigned to the particle.

As for the origin of the form, it is reasonable to assume that the form may have originated from a demonstrative, as in the case of (non-speech participant subjects in)

varieties, as well as supposedly with the 2SG form in Gweno, is acceptable in the context of a question; thus **ngi-kund-i mw-anaká shu ku** {SM1SG-love-STAT##1-woman##DEM1##NEG} 'I don't like this woman' but **ngi-kund-i mw-anaká shu fo?** 'Don't I like this woman?' (probably with surprise). I acknowledge his generous permission in sharing this and other interesting descriptive facts that should be further investigated.

Gweno, as well as in many other KB varieties, as shown in the following sections. Based on this assumption, the most plausible candidate seems to be cl. 17 **ku**, which is also used as a question word, 'where',[9] as shown in (4).

(4) **ku** as derived from locative demonstrative in Rombo-Mkuu

a. áleenda **kû**
H[10]=a-le-end-a **ku**
FOC=SM$_1$-PST.N-go-FV **where**
'Where did s/he go?'

b. alééńdá ku
a-le-enda **ku**
SM$_1$-PST.N-go-FV **NEG**
'S/he didn't go'

What is interesting, though, is that the high tone associated with the element, i.e., **ku** in Rombo-Mkuu, and that associated with coreferential markers in Gweno appear to behave in an opposite way, i.e., the high tone is realised on its underlying position when it is used as a negative marker in Gweno, which is the case in the demonstrative use in Mkuu. In contrast, the high tone shifts (and further spreads in some cases) leftward when it is used as a negative particle in Mkuu.

It should also be noted that tone is contrastive in the verb initial position as well. As illustrated in (4a), a post lexical high tone is usually realised on the initial TBU of the verb,[11] while in negative contexts, along with other types of subordinate clauses including relative clauses, the initial high tone does not occur. As pointed out by Philippson and Montlahuc (2003: 491), this high tone is associated with a preverbal focus marker[12] grammaticalised from the identificational copula (a reflex of ***ní**, the "predicative index" in Meeussen [1967: 115]), which in most cases segmentally drops. The tonal contrast between affirmative and negative clauses, conditioned by the presence and absence of the high tone marker, is widely attested in KB including in Uru, a CK language.

[9] A parallel situation is attested in Useri (E623A), another Rombo variety, where the class 17 question word **kwi** is used as a post-verbal negative marker (cf. Devos & van der Auwera 2013: 254ff).

[10] This denotes high tone accompanying with FOC.

[11] This high tone, except for several tense forms including perfect (cf. Philippson & Montlahuc 2003: 491), can be realised as a super high tone when preceded by a preverbal element which ends in a high tone.

[12] This marker, and its function, has been labelled with different names in the literature, e.g., 'stabiliser' (Nurse & Philippson 1977), 'assertion' (Dalgish 1979), 'focus' (Moshi 1988; Philippson & Montlahuc 2003), etc. See Shinagawa (2024) for a typological analysis of this morpheme in KB focusing on Rombo-Mkuu and Uru.

2.3 Uru (CK): A partial agreement system

So far we have seen two extremes of the negation marking systems in KB, namely the full agreement system in Gweno and the invariant marking system in Rombo-Mkuu. The next case from Uru can be identified as an intermediate system located between the two extremes along the scale of the complexity of main-clause negation marking.

(5) Partial agreement system in Uru

a. tʃíléolóꜜká: **pfó/ɲ̀**
tʃí-le-olok-a **pfo/ɲi**
SM_{1SG}-PST.N-fall-FV **NEG/PRON1SG**
'I did not fall'
cf. ndʒíleólo:ka
N=tʃí-le-olok-a
FOC=SM_{1SG}-PST.N-fall-FV
'I fell'

b. kuléolóꜜká: **pfó/*jo**
ku-le-olok-a **pfo/*jo**
SM_{2SG}-PST.N-fall-FV **NEG/PRON2SG**
'You did not fall'
cf. kúleólo:ka
N=ku-le-olok-a
FOC=SM_{2SG}-PST.N-fall-FV
'You fell'

c. luléolóꜜká: **pfó/*so**
lu-le-olok-a **pfo/*so**
SM_{1PL}-PST.N-fall-FV **NEG/PRON1PL**
'We did not fall'
cf. lúleólo:ka
N=lu-le-olok-a
FOC=SM_{1PL}-PST.N-fall-FV
'We fell'

d. muléolóꜜká: **pfó/*nu**
mu-le-olok-a **pfo/*nu**
SM_{2PL}-PST.N-fall-FV **NEG/PRON2PL**
'You (pl.) did not fall'
cf. múleólo:ka
N=mu-le-olok-a
FOC=SM_{2PL}-PST.N-fall-FV
'You (pl.) fell'

e. aléolóꜜká: **pfó/o**
a-le-olok-a **pfo/o**
SM_{1}-PST.N-fall-FV **NEG/DEM.M1**
'S/he did not fall'
cf. náleólo:ka
N=a-le-olok-a
FOC=SM_{1}-PST.N-fall-FV
'S/he fell'

f. waléólóꜜká: **pfó/wó**
wa-le-olok-a **pfo/wo**
SM_{2}-PST.N-fall-FV **NEG/DEM.M2**
'They did not fall'
cf. wáleólo:ka
N=wa-le-olok-a
FOC=SM_{2}-PST.N-fall-FV
'They fell'

g. kiléólóꜜká: **pfó/kjó**
ki-le-olok-a **pfo/kjo**
SM_{7}-PST.N-fall-FV **NEG/DEM.M.7**
'It (cl.7) did not fall'
cf. kíleólo:ka
N=ki-le-olok-a
FOC=SM_{7}-PST.N-fall-FV
'It (cl. 7) fell'

h. kuítʃíjó: **pfó**
ku-i-tʃí-w-a **pfo**
SM_{17}-IPFV-know-PASS-FV **NEG**(<DEM.M17)
'It (=the specific place) is not known'

In all the examples from (5a) to (5h), the invariant negative particle **pfo** appears throughout. On the other hand, the coreferential particles, which are the exact functional equivalents of those found in the Gweno system, can also be used with several regular exceptions, i.e., particles that agree with speech participant subjects are not grammatically accepted except for the case of a 1SG subject, where both the invariant **pfo** and the coreferential **ɲ** can be used. This system thus can be identified as an intermediate system between the two extremes, i.e., it partially retains the agreement system while it also accepts the invariant marker.

This 'intermediate' or 'hybrid' system of main-clause negation marking, in turn, provides several significant insights into the diachronic process of negation marking systems across KB. First, the etymological source of the invariant **pfo**, about which Devos and van der Auwera (2013: 254) raised a question, seems to become clear. If we follow the hypothesis that the original system is assumed to have been a full agreement system (cf. Nurse 1979: 274–278; Philippson & Montlahuc 2003: 496; Devos & van der Auwera 2013: 253), which is mostly retained in Gweno, and given the fact that the invariant **ku** in Mkuu supposedly originates from the locative cl. 17 demonstrative, which, in turn, is assumed to have replaced all the particles coreferential with the inflectional categories of the subject, it is feasibly likely that the invariant form in Uru, which can also be optionally used as a generalised form for all noun classes and speech participant subjects, has the same origin, i.e., cl. 17 demonstrative **pfo**,[13] which is assumed to be historically derived from ***ku-o** {PP$_{17}$-DEM.M}.

Another insight suggested by this scenario is that the variation of negation systems across KB was developed through a levelling process where the full-fledged coreferential system was first eroded by the invariant form. As evidenced in the Gweno system, it is reasonable to think that the erosion started from the 2SG marker, probably in order to avoid the synonymic clash with the cl. 1 form, as pointed out by Devos and van der Auwera (2013: 253), namely **iwe** in 2SG vs. **we** in cl. 1, by replacing the former with the class 17 locative demonstrative, which is one of the major sources of the verb-external negative particles across Bantu languages[14], with others being negation words and possessive pronouns (Devos & van der Auwera 2013: 233).[15] The next stage of the

[13] Another variant, **ko**, is attested in the demonstrative system of this language, which seems to be interchangeably used as a demonstrative (i.e., deictic marker) but cannot be used as a negative particle.

[14] However, the distribution in Eastern Bantu seems to be quite limited. According to Devos and van der Auwera (2013: 240–241), particles originating from locative demonstratives are mainly attested in western zones, namely several zone B languages, many zone H languages from the Kikongo cluster (H16), and southwestern languages from zones K, L, and R, whereas two languages from the Luyia cluster of zone JE30 are the only attestations in the east.

[15] As for the mechanism by which locatives and possessives, let alone negative words, are

process is illustrated by Uru where all but 1SG speech participants are replaced with the locative form. The final stage, where main-clause negation is consistently marked through an invariant particle, is attested in Mkuu. Table 3 is a comparative list of main-clause negation marking systems in selected KB languages, which reflects the hypothetical diachronic process.

Table 3: Agreement patterns of negation markers across varieties

Subject property		*Proto-E60?	Gweno [E65]	Uru [E622B]	Mkuu [E623C]
Speech Participants	1SG	PRON	PRON	PRON	DEM_{17}
	2SG	PRON	DEM_{17}	DEM_{17}	DEM_{17}
	1PL	PRON	PRON	DEM_{17}	DEM_{17}
	2PL	PRON	PRON	DEM_{17}	DEM_{17}
Discourse Referents	cl.1	PRON	PRON	PRON	DEM_{17}
	cl.2	PRON	PRON	PRON	DEM_{17}
	others (cl.n)	DEM_n	DEM_n	DEM_n	DEM_{17}

Given this understanding, the internal classification of negation marking systems discussed by Devos and van der Auwera (2013: 253–254) might be refined as follows. While it has been suggested that systems with the particle **pfo** or forms related to it have a different origin from the Gweno-like agreement system (Devos & van der Auwera 2013: 254), it should instead be interpreted that the presence of **pfo**, or of forms that can be traced back to locative demonstratives, is a hallmark that shows the inheritance of the original full agreement system. Along the same lines, it has been proposed that Mochi [E622A], a Central Kilimanjaro variety, inherited this agreement system based on the fact that not only **pfo**, which is identified in this context as a marker agreeing with 2SG, but also **nyi** (i.e., [ɲi]), which clearly originates from a coreferential marker agreeing with 1SG, can be used as interchangeable options. However, it may be reasonable to take the **pfo** itself as a form originating from the cl. 17 demonstrative which consistently replaces coreferential forms, as in other KB varieties, rather than taking the interchangeable use of **nyi** and **pfo**, the latter of which is ambiguously identified as a 2SG form, as evidence of

converted into negative particles, Devos and van der Auwera (2013) argue that these elements are frequently used as focus enforcers and, due to this function, these forms tend to be involved in what is known as 'Jespersen cycle', where pragmatic negative enforcers eventually take over the role of original negators as they diminish through the process of phonetic attrition. See Devos and van der Auwera (2013) for further discussions of the process.

the (partial) inheritance of the full agreement system.

2.4 Siha (WK): Utilisation of NEG2 for main-clause negation

More variation can be found in Western Kilimanjaro (WK). As illustrated in (6), Siha takes the post-initial element **ta** in addition to the invariant marker **pfo**, which is apparently cognate with the invariant **pfo** in Uru and supposedly cognate with the same form used specifically for a 2SG subject in Gweno.

(6) a. **ʃa**lek'ává **pfo**[16]
ʃa-le-k'av-a **pfo**
SM_{1SG}.**NEG-** PST.N-hit-FV **NEG**
'I did not hit'
cf. neek'ává
n-le-k'av-a
SM_{1SG}-PST.N-hit-FV
'I hit'

b. ku**ta**lék'ává **pfo**
ku-**ta-**le-k'av-a **pfo**
SM_{2SG}-**NEG-**PST.N-hit-FV NEG
'You (sg.) did not hit'
cf. kweek'ává
ku-le-k'av-a
SM_{2SG}-PST.N-hit-FV
'You (sg.) hit'

c. a**ta**lék'ává **pfo**
a-**ta-**le-k'av-a **pfo**
SM_{1}-**NEG-**PST.N-hit-FV **NEG**
'S/he did not hit'
cf. alek'ává
a-le-k'av-a
SM_{1}-PST.N-hit-FV
'S/he hit'

d. ti**ta**lék'ává **pfo**
ti-**ta-**le-k'av-a **pfo**
SM_{1PL}-**NEG-**PST.N-hit-FV **NEG**
'We did not hit'
cf. teek'ává
ti-le-k'av-a
SM_{1PL}-PST.N-hit-FV
'We hit'

e. mu**ta**lék'ává **pfo**
mu-**ta-**le-k'av-a **pfo**
SM_{2PL}-**NEG-**PST.N-hit-FV **NEG**
'You (pl.) did not hit'
cf. mweek'ává
mu-le-k'av-a
SM_{2PL}-PST.N-hit-FV
'You (pl.) hit'

f. va**ta**lék'ává **pfo**
va-**ta-**le-k'av-a **pfo**
SM_{2}-**NEG-**PST.N-hit-FV **NEG**
'They did not hit'
cf. valék'ává
va-le-k'av-a
SM_{2}-PST.N-hit-FV
'They hit'

It seems to be uncommon that the post-initial slot is utilised for main-clause negation, since it is exclusively used for negation of dependent clauses such as relative, subjunctive, and infinitive verb forms in most present-day Bantu languages (cf. Güldemann 1999:

[16] The voiceless velar stop /k/, especially in intervocalic positions, tends to be pronounced as an ejective by my consultant, though it is not clear to me whether the ejective-like pronunciation is commonly observed among other speakers.

556). This slot is also reconstructed in Proto-Bantu as a structural position for negators appearing in non-main-clause verb forms, such as ***tá-*** for infinitive and relative verb forms, and ***tį́-*** for subjunctive (Meeussen 1967: 108).

Concerning the use of tone to distinguish negative forms from corresponding affirmative forms, the tonal contrast in the verb initial position, as attested in Uru and Rombo-Mkuu, is not observed in Siha. This is apparently associated with the fact that the verb initial focus marker, **ni-** and/or its accompanying high tone, is not attested in Siha, or in other WK languages including Rwa.

2.5 Rwa (WK): Marking through the invariable particle ndi

Yet another pattern is found in Rwa, a WK variety spoken west of Siha. This language adopts a different type of clause-final negative particle, which apparently has a distinct origin from those etymologically derived from a class 17 demonstrative, such as **ku** in Mkuu and **pfo** in CK languages such as Uru as well as in some WK languages including Siha.

(7) Invariant marking through the clause-final particle **ndi** in Rwa (Shinagawa 2008: 188)

a. ṇlóliáa **ndi**
Ṇ-loli-a-a **ndi**
SM$_{1SG}$-see-FUT-FV **NEG**
'I will not see'

cf. ṇlóliáa
Ṇ-loli-a-a
SM$_{1SG}$-see-FUT-FV
'I will see'

b. tilolié **ndi**
ti-loli-ie **ndi**
SM$_{1PL}$-see-STAT **NEG**
'We do not see'

cf. tilolié
ti-loli-ie
SM$_{1PL}$-see-STAT
'We see'

c. vaworé numbé **ndi**
va-wor-ie Ø-numbé **ndi**
SM$_{2}$-hold-STAT 10-cow **NEG**
'They do not have cows'

cf. vaworé numbe
va-wor-ie Ø-numbe
SM$_{2}$-hold-STAT 10-cow
'They have cows'

As shown in (7), main-clause negation is consistently marked by the clause-final negative particle **ndi** in Rwa. Interestingly, a (quasi-)homophonous form, **nde**, is used to mark main-clause negation in the remotely related language Dawida [E74a], where, however, the form sits in the pre-initial slot of the verbal template, as shown in (8).

(8) **nde**ukúkabíeɣe
nde=u-ku-kab-ieɣe
NEG=SM_1-OM_{2SG}-hit-PST
'S/he did not hit you'

According to Kamba-Muzenga (1978: 99), this form, in turn, is classified into a group of cognate forms distributed in neighbouring languages, notably including Central Kenyan languages [E50], as shown in (9). All of these forms are historically traced back to ***tíɪ̩** (> **ndi** > **ndV**), which is reconstructed as the negative predicative index 'it is not' (Meeussen 1967: 115).[17]

(9) Pre-initial negative markers historically relatable to ***tíɪ̩** in selected Zone E languages (Kamba-Muzenga [1978: 99])
a. Taita/Dawida [E74] **ndi**dikabagha
'nous ne frappons pas (We don't hit)'
b. Gikuyu [E51] **ndu**gwataga
'tu ne tiens pas (You don't hold)'
c. Kamba [E55] **ndu**neekoota
'tu n'as pas tiré (You didn't shoot)'

The homonymity of these forms may suggest that the form **ndi** in Rwa is etymologically identified as a cognate with the prefixal **nde** in Dawida, which is shared with neighbouring E50 languages. If this is the case, then this also suggests that the form may have been introduced through transposition so that it could fit in the morphosyntactic template in KB, where the negative particle consistently appears in the clause-final position.

3. Summary with additional notes for further investigation

The main-clause negation systems in selected KB languages that are described in this article can be summarised as in Table 4 (additional information on Mochi is from Nurse [1979: 275], as cited in Devos & van der Auwera [2013: 253]).

[17] Devos and van der Auwera (2013: 254ff) also refer to the possibility that the **ndi** in Rwa/Meru [E621A] may not be locative in origin but may be related to a rhetorical question marker (based on pers. comm. with Gérard Philippson). Though it is clearly beyond the scope of this paper, it appears to be plausible that the negative predicative index, or functionally saying, a negative copula, can be used as a tag question ('isn't it?'), which in turn might be pragmatically used as a 'rhetorical question marker'. More information about the practical usage of this element is needed to clarify this point.

Table 4: Summary of the main-clause negation marking across KB[18]

		Verb-external			Verb-internal	
		Agreement	Form	Source	Form	Source
Gweno [E65]		Full	(various)	PRON (SP), DEM (NC)		
CK	Mochi [E622A]	Partial?	**nyi**[19] **pfo** (others?)	PRON.1SG cl. 17 DEM		
	Uru [E622B]	Partial	**nyi** **pfo** (various)	PRON.1SG cl. 17 DEM DEM (NC)		
Rombo	Mkuu [E623C]	Invariant	**ku (pfo)**	cl. 17 DEM		
WK	Siha [E621C]	Invariant	**pfo**	cl. 17 DEM	**ta**	NEG2 in PB
	Rwa [E621A]	Invariant	**ndi**	***tí̩** 'NEG index'		
Dawida [E74a]		Invariant	**nde=**	***tí̩** 'NEG index'		

As described in 2.1 and 2.2, the Gweno system and the Mkuu system illustrate two extremes in terms of the structural complexity observed in the contemporary languages within the scope of this study. Intermediate systems are observed in CK, including the hybrid system in Uru where the invariant form **pfo** is pervasively used for all subjects, while agreement markers can also be optionally used only for non-speech participant subjects. On the other hand, the two invariant marking systems seem to be contrastive at least from a diachronic perspective, i.e., the Siha system shows a clear parallel with the Mkuu system, which thus can be placed at the simplest extreme, while the Rwa system

[18] SP: Speech participants; NC: Noun classes.

[19] It is interesting to note that, as pointed out in Devos and van der Auwera (2013: 253), the PRON.1SG **nyi** in Mochi appears to have functioned as a generalised negative particle, unlike its cognate in Uru, where the form seems to be specifically used with a 1SG subject. More importantly, especially in the case of Mochi, where both the 1SG form and 2SG form **pfo** have been grammaticalised as across-the-board negative markers, it may be plausible to think that that these forms might have been associated with pragmatic discourse functions. For example, PRON.1SG might convey the speaker's emphatic confirmation of the negated proposition (e.g., 'I'm sure [it's not...]'), while PRON.2SG might suggest a request for the hearer's confirmation (e.g. '[It's not...], don't you think?'). If so, such interlocutory pragmatic functions might have been a key factor motivating the expansion of these forms into generalised negative markers. I am grateful to Benji Wald for bringing this point to my attention.

seems to have a different historical origin, and thus cannot be directly located in the diachronic scenario of diversification, where the original system has been 'eroded' by the forms originating from cl. 17 locative demonstratives.

It should also be noted that this scenario is based on the assumption that the full agreement system is the one that is reconstructed as the proto system as suggested in the literature (cf. Nurse [1979: 274–278]; Philippson & Montlahuc [2003: 496]). This assumption, however, may also need further verification, since the factors that motivate the invariant form to replace agreement forms, first those agreeing with speech participants and then with all non-speech participants, have not been completely clarified.

On the other hand, at least within the range of languages described in this paper, only Rwa seems to have inherited another origin, which is shared with Dawida, as well as with other E50 languages spoken northeast of the Kilimanjaro area. This may suggest that the agreement system might have been a later innovation in Kilimanjaro, which replaced the old consistent marking system with the forms that historically trace back to the negative predicative index ***tí̩**. These diachronic assumptions are no doubt tentative working hypotheses that need substantial verification through extensive data not covered in this study.

Abbreviations

1, 2, 3...	noun classes (when referred to as agreement properties, they are subscripted, e.g., SM_1, OM_2 etc...)	OM	object marker
1SG, 2PL etc.	person and number	PASS	passive
DEM	demonstrative	POSS	possessive
DEM.M	demonstrative (medial)	PRON	pronoun
DEM.P	demonstrative (proximate)	PRF	perfect
FOC	focus/main clause marker	PRS	present
FUT	future	PST	past
FUT.R	remote future	PST.N	near past
FV	final vowel (default inflectional suffix)	PST.R	remote past
HAB	habitual	SM	subject marker
IPFV	imperfective	STAT	stative
INF	infinitive	-	affix boundary
NEG	negation	=	clitic boundary
NOT.YET	'not yet tense', as mentioned in Philippson & Nurse (2000)	##	word boundary

References

Dalgish, Gerard M. 1979. The syntax and semantics of the morpheme *ni* in KiVunjo (Chaga). *Studies in African Linguistics* 10(1): 47–63.

Devos, Maud, and Johan van der Auwera. 2013. Jespersen cycles in Bantu: Double and triple negation. *Journal of African Languages and Linguistics* 34(2): 205–274.

Güldemann, Tom. 1999. The genesis of verbal negation in Bantu and its dependency on functional features of clause types. In Hombert, Jean-Marie, and Larry M. Hyman (eds.) *Bantu historical linguistics*. California: CSLI, pp. 545–587.

Güldemann, Tom. 2022. Predicate structure and argument indexing in early Bantu. In Koen Bostoen, Gilles-Maurice de Schryver, Rozenn Guérois, and Sara Pacchiarotti (eds.) *On reconstructing Proto-Bantu grammar*. Berlin: Language Science Press, pp. 387–421. DOI: 10.5281/zenodo.7575831

Guthrie, Malcolm. 1967–71. *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. Vols. 1–4. Farnborough: Gregg Press.

Hammarström, Harald. 2019. An inventory of Bantu languages. In Van de Velde, Mark, Koen Bostoen, Derek Nurse, and Gérard Philippson (eds.) *The Bantu languages. Second edition*. London: Routledge, pp. 17–78.

Kamba-Muzenga, Jean Georges. 1978. *Les formes négatives dans les langues bantoues,* Tervuren: Koninklijk Musuem voor Midden-Afrika.

Maho, Jouni F. 2009. NUGL Online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages. Available online at https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf

Meeussen, Achille E. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3: 80–122.

Moshi, Lioba. 1988. A functional typology of "ni" in Kivunjo (Chaga). *Studies in the Linguistic Sciences* 18(1): 105–113.

Nurse, Derek. 1979. *Classification of the Chaga dialects*. Hamburg: Helmut Buske Verlag.

Nurse, Derek 1981. Chaga/Taita. In Hinnebusch, Tom, Derek Nurse, and Martin Mould (eds.) *Studies in the classification of Eastern Bantu languages* (Sprache und Geschichte in Afrika, Beiheft 3). Hamburg: Helmut Buske Verlag, pp. 127–180.

Nurse, Derek. 2008. *Tense and aspect in Bantu*. Oxford: Oxford University Press.

Nurse, Derek, and Gérard Philippson. 1977. Tones in Old Moshi. *Studies in African Linguistics* 8(1): 49–80.

Philippson, Gérard, and Marie-Laure Montlahuc. 2003. Kilimanjaro Bantu (E60 and E74). In Nurse, Derek, and Gérard Philippson (eds.) *The Bantu languages*. Oxon: Routledge, pp. 475–500.

Philippson, Gérard, and Derek Nurse. 2000. Gweno, a little known Bantu Language of

Northern Tanzania. In Kahigi, K. K., Yared M. Kihore, and Maarten Mous (eds.) *Lugha za Tanzania/Languages of Tanzania: A study dedicated to the memory of the late Prof. C. Maganga*. Leiden: CNWS, pp. 233–284.

Shinagawa, Daisuke. 2008. Verbal morphology of Rwa (Bantu E61): A descriptive study [in Japanese]. Doctoral thesis, Nagoya University.

Shinagawa, Daisuke. 2015. *A grammatical sketch of Chaga-Rombo* [in Japanese]. Tokyo: ILCAA.

Shinagawa, Daisuke. 2024. A micro-parametric approach to focus marking *ni* in Kilimanjaro Bantu languages: With special reference to Rombo-Mkuu and Uru. In Bloom-Strom, Eva-Marie, Hannah Gibson, Rozenn Guérois, and Lutz Marten (eds.) *Morphosyntactic variation in Bantu*. Oxford: Oxford University Press.

Shinagawa, Daisuke, and Lutz Marten. 2021. Micro-typological covariation of negation and focus marking morphology in Bantu languages. *Gengo Kenkyu* [Journal of the Linguistic Society of Japan] 160: 215–248.

ISBN 978-4-86337-545-1

Center for Language Studies, Kyoto Sangyo University

Research Institute for Languages and Cultures of Asia and Africa,
Tokyo University of Foreign Studies