FACULTY OF ARTS AND PHILOSOPHY

18.02.26, LLACAN SEMINARS, CNRS-LLACAN

VERB STEM VARIATION IN CAMEROONIAN BANTOID

Elisabeth J. Kerr

GHENT UNIVERSITY

BantUGent

fwo

Langage, Langues et Cultures d'Afrique
Llacan

TALK OUTLINE

INTRODUCTION

GHENT
UNIVERSITY

THE VERB STEM

- **Verb stem** = verb root (aka *radical*) + suffixes (aka *extensions*)
- **Verbal word** = prefixes + verb stem
- In Bantu/Bantoid, verbal suffixes typically derivational and prefixes typically inflectional
- High number of co-occurring verbal affixes means Bantu languages are considered highly synthetic/agglutinative (e.g. Nurse 2008; Good 2012

(1) **Ndiɗaɓarirɣa** paDú:nda.
/ndi-ɗa-ɓar-ir-ɣ-a pa-Dunda/
SM.1SG-PFV-give.birth-APPL-PASS-FV 16-Dunda
'I was born in Dunda.' (lit. 'I was given birth to in Dunda.')
(Barwe (N45, Mozambique); Zahran et al. to appear:§3.1, adapted)

(2) **ddiyááguliyedhíwa** naámbédde álêddo.
/ddi-hí-á-gul-ih-edh-íw-a naámbédde álêddo/
SM.1SG-PFV.DJ-OM.2-eat-CAUS-APPL-PASS-FV 1a.maize 2.guest
'I made the maize be sold to the guests.'
(Cuwabo (P34, Mozambique); Guérois 2015:261, glosses adapted)

(3) **Ndiɗaɓariɾya** paDú:nda.
/ndi-ɗa-ɓar-iɾ-ɣ-a pa-Dunda/
SM.1SG-PFV-give.birth-APPL-PASS-FV 16-Dunda
'I was born in Dunda.' (lit. 'I was given birth to in Dunda.')
(Barwe (N45, Mozambique); Zahran et al. to appear:§3.1, adapted)

(4) **ddiyááguliyedhíwa** naámbédde álêddo.
/ddi-hí-á-gul-ih-edh-íw-a naámbédde álêddo/
SM.1SG-PFV.DJ-OM.2-eat-CAUS-APPL-PASS-FV 1a.maize 2.guest
'I made the maize be sold to the guests.'
(Cuwabo (P34, Mozambique); Guérois 2015:261, glosses adapted)

VERB STEM COMPLEXITY IN BANTU

PROTO-BANTU VERB TEMPLATE

Prefix aka pre-stem cluster					Stem cluster		
-4	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**
(pre-initial)	initial	(post-initial)$^{+}$	(pre-radical)	radical	(pre-final)$^{+}$	final	(post-final)
TAMP clause type	subject	TAMP clause type	object	verb	derivation TAMP	TAMP	participant P clause type

Table: Proto-Bantu verb schema (Meeussen 1967), via Güldemann 2022:388. (...) = optional material; $^{+}$ = can recur; T = tense; A = aspect; M = mood; P = polarity.

VERB STEM COMPLEXITY IN NON-BANTU BANTOID

"Many Bantoid languages have few extensions, often limited to one per verb root. In addition, although Bantu languages typically allow more than one extension in sequence, many of the Bantoid languages allow only one extension per verb root."

(Hyman 2018:179)

BANTU VS BANTOID

Table 10: Canonical Bantu contrasted with Bantoid

	Canonical Bantu	Bantoid
phonology	minimum word = 2 syllables	maximum stem = 2 or 3 syllables
morphology	highly synthetic, agglutinative	less so, gradual move towards analyticity
unmarked objects	multiple	one per verb
marked objects	head marking on verb	prepositions, serial verbs

(Hyman 2018:192)

DIACHRONY

- Hyman (2004, 2017, 2018): syntheticity in Bantu is old; Bantoid shorter forms are cases of morphosyntactic change driven by areal phonological constraints which led to reduction of verb stems

- Güldemann (2022): syntheticity in Bantu is new; Bantoid shorter forms are the older stage, which became synthetic in Narrow Bantu via univerbation

DIACHRONY

- FWO postdoc project (MBAM project, Oct 2025 - Sep 2027): Use Bantu/Bantoid languages of Cameroon as test case for modelling synthetic ↔ analytic change

- Today: Consideration of variation in stem length + degree of stem-initial prominence effects

CAMEROONIAN BANTOID

Figure: Map of Niger-Congo subfamilies (Good 2017:473)

WESTERN MBAM: TUNEN, NYOKON, NOMAANDÉ, ATOMB

Figure: Map showing location of Western Mbam languages Tunen [tvu] (Bantu A44), Nyokon [nvo] (Bantu A45), Nomaandé [lem] (Bantu A46), and Atomb [ttf] (Bantu A461) within Cameroon

MAXIMALITY CONSTRAINTS

GHENT
UNIVERSITY

PHONOLOGY AS DRIVER OF MORPHOSYNTACTIC CHANGE?

- Hyman (2004, 2017, 2018): syntheticity in Bantu is old; Bantoid shorter forms are cases of morphosyntactic change **driven by areal phonological constraints**

MAXIMALITY CONSTRAINTS

- Cameroonian Bantoid languages show shorter verb roots & limited suffixal expansion

vì 'come'	*vì-sɔ́* 'bring near'
fɛ́n 'be black'	*fɛ́n-sɔ́* 'make black'
dhú 'go'	*dhú-sɔ́* 'carry away'
zhɨ́ 'eat'	*zhɨ́-sɔ́* 'feed'
búŋ 'melt'	*búŋ-sɔ́* 'cause to melt'
lyɔ́m 'hurt self'	*lyɔ́m-sɔ́* 'hurt s.o.'
vì 'come'	*vì-nɔ́* 'come with'
tsí 'spend night'	*tsí-nɔ́* '... with a woman'
cò 'pass'	*cò-mɔ́* 'meet and pass'
kwèʔè 'think'	*kwèʔ-mɔ́* 'think together'
gè 'share'	*gè-mɔ́* 'share equally'
táŋ 'count'	*táŋ-mɔ́* 'quarrel'
shɨ̀ʔ 'measure'	*shɨ̀ʔ-mɔ́* 'compare measures'

Figure: Babanki (Bantoid) verb forms listed by Hyman (2018:12).

MAXIMALITY CONSTRAINTS

VARIATION WITHIN CAMEROONIAN BANTOID

Table 4. Features of grammatical tone in NW Bantu verb phrases

Language		Phonemic tone inventory	Syllable max. of verb stem	Verb radical tone	GT change		Phrase-medial	GT on IAV
Name	Zone				STAMP	Verb stem	GT change	
Balon	(A13)	H, L	3	L vs. H	yes	yes	yes	no
Bakweri	(A22)	H, L, ^{i}H	4	L vs. H	yes	yes	yes	no
Yasa	(A33a)	H, L	4	L vs. H	yes	yes	yes	no
Basaa	(A43)	H, L	3	L vs. H	yes	yes	yes	yes
Bafia	(A53)	H, L, M	3	(L vs. H)	yes	yes	yes	yes
Nuasue	(A62A)	H, L, LH, ∅	4	L vs. H	yes	yes	no	no
Eton	(A71)	H, L, D	3	L vs. H	yes	yes	yes	yes
Gyeli	(A801)	H, L, HL, LH, ∅	3, mostly 2	L vs. H	yes	yes	yes	yes
Kwakum	(A91)	H, L, ∅	3, mostly 2	L vs. non-L	yes	yes	yes	no
Orungu	(B11b)	H, L, ^{i}H, HL, H^{i}H	4	no contrast	yes	yes	no	no
Kota	(B25)	H, L	4	L vs. H	yes	yes	no	no
Pove	(B305)	H, L, ^{i}H	3	L vs. H	yes	yes	yes	no

Figure: Overview of Northwestern Bantu languages showing maximality constraints (“Syllable max. of verb stem”) max of verb stems (Grimm 2025:17)

WEAKER STEM LENGTH RESTRICTIONS IN MBAM BANTU

- Tunen and Nomaandé allow multiple suffixes; verb stems can reach 4-5 syllables

(5) Mɛ ná nɛmbɛa **bésálél.**
/Mɛ ná nɛ-mbɛa **bés-Vl-Vl**/
SM.1SG PST2 5-melon cut-DIM-DIM
'I cut a tiny little bit of the melon.' (Tunen (A44, Cameroon); own data)

(6) Muə́kúnyi mú mə **tə́límínyisi** ɔ́ɔ́n.
/mɔ-ə́kúnyi mú ma **tə́límíny-əsi** ɔ-ɔ́nɔ/
1-teacher SM.1 PST stand-CAUS 1-child
`The teacher made the child stand up.' (Nomaandé (A46, Cameroon); own data)

MAXIMALITY CONSTRAINTS: SUMMARY

- Maximality constraints = phonological restriction on size of verb stem
- Variation in Cameroonian Bantoid languages in the degree of maximality constraints: 2, 3, 4, 5+ syllable verb stems
- Mbam Bantu languages Tunen and Nomaandé have weaker maximality restrictions

STEM-INITIAL PROMINENCE

GHENT
UNIVERSITY

STEM-INITIAL PROMINENCE

- Maximality constraints have been linked to **stem-initial prominence** (Idiatov and Van de Velde 2016; Lionnet and Hyman 2018; Hyman et al. 2019)

  Stem-initial prominence (SIP)
  Prosodic prominence of the beginning of the verb stem.

- SIP may be manifested by phoneme inventory, phonetic effects (e.g. lengthening of C_1/V_1 material), and position-based phonological rules (e.g. lenition, tone) (Lionnet and Hyman 2018; Hyman et al. 2019; Lovegren 2012, i.a.)

SIP

Phoneme inventory

- Lionnet (2017): overview of strong vs weak SIP based on skew of phoneme inventory
- Argues that SIP is an areal feature, mainly but not exclusively found in Niger-Congo

Example of a very strong SIP language: Lua (Adamawa, Bua) [Boyeldieu 1985]

❖C_1V_1 | (C).(CV)(C)

C1	p	t	c	k	ʔ	b	d	j	g	mb	nd	nj	ŋg	ɓ	ɗ	s	h	m	n	ɲ		l	r	w	w̃	y
C2+	/B, D, J, G/															s		m	n	ɲ	ŋ	l	r	w		y

V1	i	e	ɛ [ia]	ɨ	ə	a	u	o	ɔ [ua]	ĩ	ẽ	ɛ̃ [iã]	ə̃	ã	ũ	õ	ɔ̃ [ũã]	+ length
V2	i	e		ɨ	ə	a	u	o										no length

C1 > C2+	V1 > V2+	Stem type	Stem max
strong	strong	All	2σ

Example of a weaker SIP language: Gbaya [Moñino & Roulon 1972]

❖C_1 | V_1(C).(CV)(C)

C1	p	t	k	kp	ʔ	b	d	g	gb	mb	nd	ŋg	ŋgb	ɓ	ɗ	f	s	h	v	z	m	n	ɲ	ŋ	ŋm	l	r	w	y
C2+		t	k	kp						mb	nd	ŋg	ŋgb	ɓ	ɗ	f	s			z	m	n	ɲ	ŋ	ŋm	l	r	w	y	v
C/_#	p	t		k																	m	n		ŋ		l	r			

V1	i	e	ɛ	a	ɔ	o	u	ĩ	ẽ	ã	ɔ̃	ũ
V2	i	e	ɛ	a	ɔ	o	u	ĩ	ẽ	ã	ɔ̃	ũ

C1 > C2+	V1 > V2+	Stem type	Stem max
weak	(no)	All	2σ (3σ)

Figure: SIP in Lua and Gbaya (Boyeldieu 1985 and Monino & Roulon 1972, via Lionnet 2017).

- Phonological prominence of stem-initial position can be reflected by phonetic/acoustic effects, e.g. lengthening, intensity
  - Such phonetic effects interact with phoneme inventory - e.g. allowing for consonants in C1 position that require higher articulatory effort (e.g. labiovelars)

- Phonetic correlates less well-studied than phoneme inventories

IDIATOV & VAN DE VELDE (2016)

- Idiatov and Van de Velde (2016): lengthening of C_1 found in 7 of 9 Cameroonian Bantoid languages
  - (Eton (A71/[eto], Cam), Bafut (Bantoid/[bfd], Cam), Basaa (A43a/[bas], Cameroon), Okak Fang (A75/[fan], Gabon), Kota (B25/[koq], Gabon), Kwasio (A81/[nmg], Cam), and Bagyeli (A81/[gyi], Cam))
- No increased C_1 duration for Bapuku (A30/[bng], Equatorial Guinea/Gabon) or Orungu (B11b/[mye], Gabon)
- Increased C_1 also found for Ejagham (Ekoid/[etu], Cam), Gbaya Mbodomo (Gbaya/[gmm], Cam), Vute (Mambiloid/[vut], Cam), and Yemba (Bamileke/[ybb], Cam) (Dmitry Idiatov, p.c.)

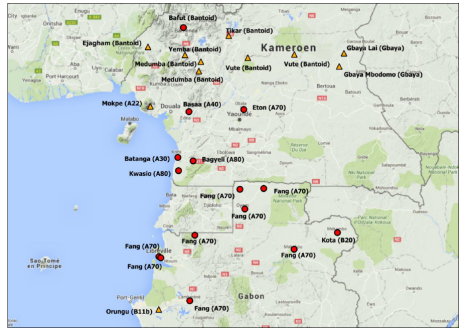


Figure: Map of languages in Idiatov and Van de Velde (2016)'s study (Idiatov and Van de Velde 2016:8).

IDIATOV & VAN DE VELDE (2016): C1 LENGTH

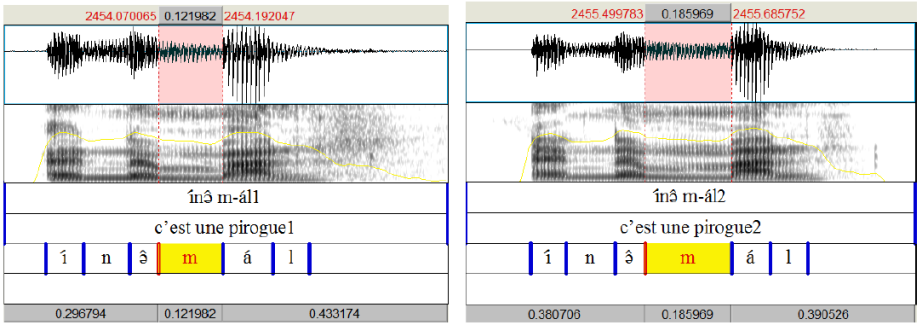
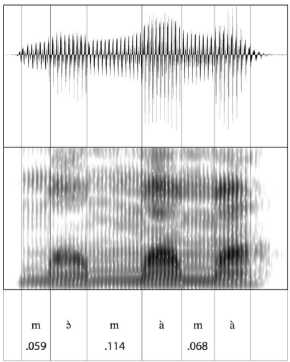


Figure: Example length of /m/ in C1 position in Eton (Idiatov and Van de Velde 2016:14)

IDIATOV & VAN DE VELDE (2016): C1 LENGTH

- Idiatov and Van de Velde (2016) report application of stem-initial prominence to nonce words
- e.g. Eton nonce word *mə-mama*:



Pfx = 0.059ms

C1 = 0.114ms

C2 = 0.68ms

Figure: Annotated Eton nonce word (Idiatov and Van de Velde 2016:3).

REPLICATION STUDY

- Replication study conducted for Tunen (Bantu A44 [tvu]) and Nomaandé (Bantu A46 [lem])

RQ: Do Tunen and Nomaandé show lengthening of C_1?

→ Test for /m/ and /b-p/ using replication of Idiatov and Van de Velde (2016)'s method

METHODOLOGY

- Fieldwork experiment conducted in-situ with 4 speakers of Tunen (2 M, 2 F) and 4 speakers of Nomaandé (3 M, 1 F)
- Drafted wordlists controlling for /m/ and /b-p/ across prefixal, C_1, C_2, and C_3 contexts
- Checked wordlists in advance with speakers and checked sentential frame
- Sessions recorded in a quiet environment (Zoom H5 + Shure WH20XLR headphone microphone + video)
- Interviews conducted in French; discourse contexts then provided by the elicitor during the experiment in the target language

METHODOLOGY: WORDLISTS

Item	Cat.	Meaning	Condition(s)
màmá	N	'mama'	m_C_1, m_C_2
bamámá	N	'mamas'	b_Pfx, m_C_1, m_C_2
àkʒ̀	V	'to do'	filler
mahʊ̀ma	N	'forests'	m_Pfx, m_C_2
ɛ̀nɛ̀na	N	'neck'	m_C_1
mɛnyama	N	'animals'	m_Pfx, m_C_2
tʊ́tʊ́mámá túnúɔ	N+Mod	'all the fruits'	m_C_2, m_C_{2Redup}
mísí mìkímɔ	N+Mod	'the whole world'	m_C_1
mɔ́nɔ́	N	'child'	m_C_1
mɔnɔ́ mɛ̀lɛ̀mɛndɔ	N_Comp	'boy'	m_C_1
mɔnɔ́ ámɔtɛ	N+Mod	'one child'	m_C_1
wɔmɔtɛ́ mɔnɔ́	D+N	'a certain child'	m_C_1
hìwɔ	N	'fire'	filler
mɔnífɔ	N	'water'	m_Pfx
íbʊ̀ɔ	N	'salt'	b_C_1
tʊnɔnɨ	N	'Tunen'	filler
bɛfáfálaka	N	'butterflies'	b_Pfx
ɛmalamandɛ	N	'thunder' (sg.)	m_C_1, m_C_3
mamalamandɛ	N	'thunder' (pl.)	m_Pfx, m_C_1, m_C_3
ámanâ bùlìbɔ	Other	'only tomorrow'; 'see you tomorrow'	m_C_1, b_C_1
ɛmanya	N	'intelligence'	m_C_1
ɛ́nàsɛ	Other	'very soon'	filler
mambɔmɔ	N	'forests'	m_Pfx, m_C_2
mambɔmɔ má Kamɛ́rúnɑ	N	'forests of Cameroon'	m_Pfx, m_C_2, m_C_2
bɛlama	N	'vegetables'	b_Pfx, m_C_2
ámɛ́la	V	'be pregnant'	m_C_1
ɔmɛma	V	'to swallow'	m_C_1
mɛma!	V	'swallow!'	m_C_1
tʊmɔmɔ	N	'fingers'	filler
mamɛna	N	'necks'	m_Pfx, m_C_1
tʃfátʃfá	Other	'quickly'	filler
ɔ́fáma	V	'to arrive'	m_C_2
ɔbʊ́ŋɔ	V	'to find'	b_C_1
ɔmanya	V	'to know'	m_C_1
bíʃʊ́ákâ	V	'to wash oneself'	b_Pfx
usʊ́bɔ	V	'to drum'	b_C_2
ɛ́hɔka	N	'place'	b_C_1
ɔbâ	V	'to be'	b_C_1
bábɛ̀!	V	'turn! (couscous)'	b_C_2
uhʊ́bɔ	V	'to turn (couscous)'	b_C_2
bɛlama bɛ̀ Kalɛ́bɛ	N+Mod	'Caleb's vegetables'	b_Pfx, b_C_1, b_C_3
mamama máfándɛ	N+Mod	'two *mamamas*'	m_Pfx, m_C_1, m_C_2
mɛmama máfándɛ	N+Mod	'two *mɛmamas*'	m_Pfx, m_C_1, m_C_2
talɔkɔ	N/A	N/A	filler
bababa bálálɔ	N+Mod	'three *bababas*'	b_Pfx, b_C_1, b_C_2
babɔbɔ bálálɔ	N+Mod	'three *babɔbɔs*'	b_Pfx, b_C_1, b_C_2

Item	Cat.	Meaning	Condition(s)
1. mɔaɔmɔ	N	mama	m_C_1, m_C_2
2. amɛna	N	neck	m_C_1
3. nyɛbána	N	breast	b_C_1
4. hínonyí	N	birds	filler
5. ɔmúá mɔnyífɔ	V	to drink water	m_C_2, m_Pfx
6. makɛ̂lɛ́	N	salt	m_Pfx
7. bʊúsə	N	face	b
8. ɛnyama	N	animal	m_C_2
9. íbʊ̀nyi ínyísə	N+Mod	four goats	b_C_1
10. mabala	N	urine (pl.)	m_Pfx, b_C_1
11. ɛŋkɛ̀ɛ́mùaná	N	sky	m_C_2
12. yɔɔbɔ́	N	skin	b_C_2
13. ɔɔbɔ́	N	?	
14. ɔhámə	V	to go out	m_C_2
15. owəəbə	V	to steal	b_C_2
16. nəbɛ́mbɛ	N	tongue	b_C_1
17. mɛnâ mɛnyífɔ!	V	drink water!	m_C_2, m_Pfx
18. əətú	N	ears	filler
19. ɔ́ɔ́cɔ́	N	sun	filler
20. ɔɔcɔ́	N	fire	filler
21. ɔɔcɔ	N	man	filler
22. bʊnɔ́ɔ́ búmɔtɛ́	N+Mod	one village	b_Pfx
23. háma!	V	go out!	m_C_2
24. obíkútə	V	to shave oneself	b_Pfx
25. ɔfaaya	V	to come	filler
26. əbʊ́ŋɔ	V	to find	b_C_1
27. búmâkə bíbábí!	V	chase the he-goats!	b_Pfx, b_C_1 (x2), b_C_2
28. malaka má Kalɛ́bɛ	N+Mod	Caleb's clothes	m_Pfx, m_C_1
29. bɛnyéma bɛ̀ Bálána	N+Mod	Miriam's things	b_Pfx, b_C_1
30. mamama máfándí	N	two *mamamas*	m_Pfx, m_C_1, m_C_2
31. nyɛmama nyémɔtɛ́	N	one *nyɛmama*	m_C_1, m_C_2
32. talɔkɔ	N/A	N/A	filler
33. bababa bátátɔ	N	three *bababas*	b_Pfx, b_C_1 b_C_2
34. babɔbɔ bátátɔ	N	three *babɔbɔs*	b_Pfx, b_C_1 b_C_2

Figure: Tunen and Nomaandé real and nonce words.

METHODOLOGY: ISOLATION/FRAME

- Part I: SIP Words
  - Display each word on screen; speaker pronounces x1 in isolation and x3 in context 'I say X again'
  - Mix of /m/ and /p-b/ items and fillers

Figure: Tunen consultant performing Part I.

METHODOLOGY: NONCE WORDS

- Part II: SIP Nonce Words
  - As with Part I, but with nonce words that obey the language's phonotactics
  - Nonce words given with numeral modifiers to cue N status

Figure: Example Nomaandé nonce word *babɔbɔ* cued as N with structure *ba-bɔbɔ* by co-occurrence with numeral *bátátɔ́* 'three').

METHODOLOGY: DISCOURSE CONTEXT

- Part III: SIP Context
  - Dialogic format with elicitor; background, new information focus, corrective focus, correction of elicitor's mispronunciation (pseudo-argument format)

Figure: Corrective focus context between EK and consultant.

METHODOLOGY: DISCOURSE CONTEXT

- Part III: SIP Context
  - Dialogic format with elicitor; background (Bg), new information focus (FocIF), corrective focus (FocCorW), correction of mispronunciation (FocContrPfx, FocContrC1, FocContrC2)

(7) a. ɔ́ ndɔ **ibusə́** sinə eé?
SM.2SG PRS 9.salt see Q
'Do you see salt?'

b. bɔ́ɔ, **mamalamanda** á mɛ́ ndɔ sin.
no 6.thunder COP SM.1SG.REL PRS see
'No, I see [thunder]$_{FOC}$.' (Tunen)

METHODOLOGY

- Part III: SIP Context
- Dialogic format with elicitor; background (Bg), new information focus (FocIF), corrective focus (FocCorW), correction of mispronunciation (FocContrPfx, FocContrC1, FocContrC2)

(8) a. buhúnyi ɛ́cɛ nakɛ́lɛ́.
14.word COP nakɛ́lɛ́
'The word is nakɛ́lɛ́.'

b. əə?əə, buhúnyi ɛ́cɛ **makɛ́lɛ́**, tícáŋa **nakɛ́lɛ́.**
no 14.word COP 6.salt COP.NEG nakɛ́lɛ́
'No, the word is ***m*akɛ́lɛ́**, not ***n*akɛ́lɛ́.**' (Nomaandé)

RECORDING FILES

- Resulting files for Part I (SIPIsolation) + Part II (SIPNonceWords) + Part III (SIPContext):

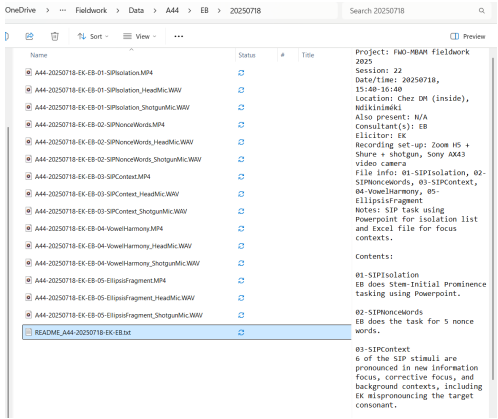


Figure: Example set of recording files and metadata.

RECORDINGS → TEXTGRID ANNOTATIONS

- Head mic .WAV files opened in Praat and annotated to TextGrid

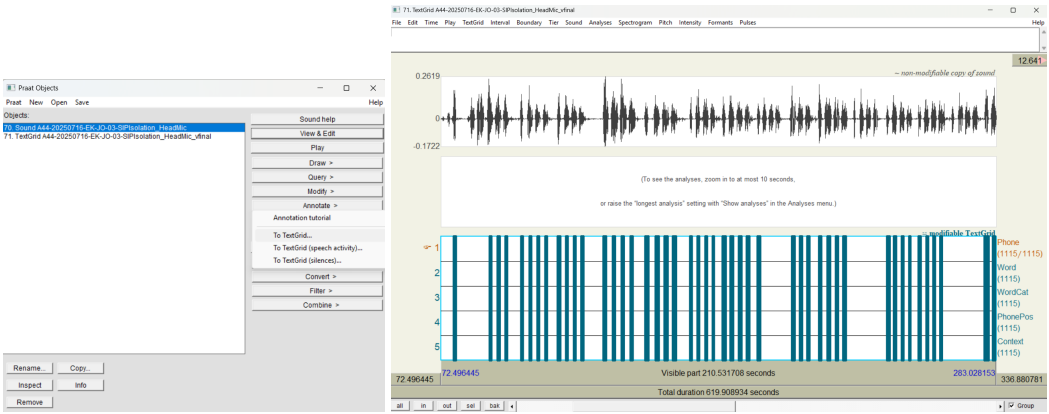


Figure: Example annotated TextGrid in Praat.

RECORDINGS ⟶ TEXTGRID ANNOTATIONS

- Head mic .WAV files opened in Praat and annotated to TextGrid

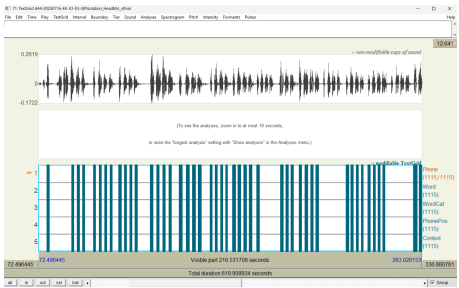


Figure: Example annotated TextGrid in Praat.

⇒ 5 annotation tiers:

1. **Phone** (m, b, á, ə, ŋ, etc.)
2. **Word** (ɛmanya, ibusə́, etc.)
3. **WordCat** (N, NComp, NNonce, NMispronun, NProp; V; Assoc, Mod, Other)
4. **PhonePos** (Pfx, C1, C2, C3, C4; V1, V2, V3, V4; C1Redup, C2Redup, V1Redup, V2Redup; NA)
5. **Context** (Iso, Fr, FrIso; Bg; FocIF; FocCorW; FocContrPfx, FocContrC1, ForContrC2)

ANNOTATION DECISIONS

- For consultants that repeated an item more than 3 times, take 3 best-quality audio; if all good quality, take first 3
- Annotate manually in Praat based on waveform + spectrogram (showing formants and pitch trace)
- Annotate surface tone on Phone tier and use lexical tone for Word tier
  - e.g. /ibusə́/ 'salt' pronounced in Iso context as [ibusə] (LLL): Phone = ə; Word = ibusə́
- Annotate entire stop (closure to release, not showing burst location)
- Annotate boundaries at midpoint of formant transition

ANNOTATION DECISIONS

VOWEL LENGTH

- Include or exclude tail of vowels?

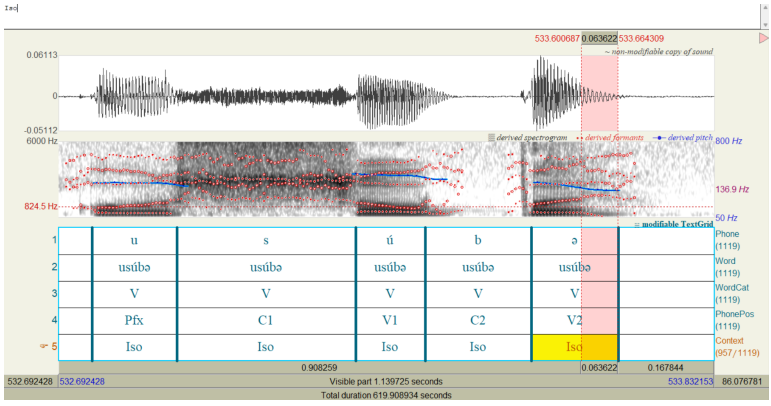
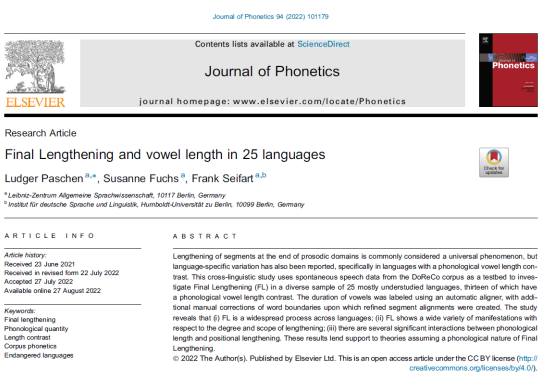


Figure: Example annotation of vowel in final position including tail.

ANNOTATION DECISIONS

VOWEL LENGTH

- Included vowel tails, but therefore exclude vowels in final position from analysis
  - Exclude items in Iso context; use sentential frame and discourse contexts
- Lengthening of final vowels reported cross-linguistically; likely independent of SIP (see e.g. Seifart et al. 2021; Paschen et al. 2022)



Journal of Phonetics 94 (2022) 101179

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Research Article

Final Lengthening and vowel length in 25 languages

Ludger Paschen [a,*], Susanne Fuchs [a], Frank Seifart [a,b]

[a] Leibniz-Zentrum Allgemeine Sprachwissenschaft, 10117 Berlin, Germany
[b] Institut für deutsche Sprache und Linguistik, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

ARTICLE INFO

Article history:
Received 23 June 2021
Received in revised form 22 July 2022
Accepted 27 July 2022
Available online 27 August 2022

Keywords:
Final lengthening
Phonological quantity
Length contrast
Corpus phonetics
Endangered languages

ABSTRACT

Lengthening of segments at the end of prosodic domains is commonly considered a universal phenomenon, but language-specific variation has also been reported, specifically in languages with a phonological vowel length contrast. This cross-linguistic study uses spontaneous speech data from the DoReCo corpus as a testbed to investigate Final Lengthening (FL) in a diverse sample of 25 mostly understudied languages, thirteen of which have a phonological vowel length contrast. The duration of vowels was labeled using an automatic aligner, with additional manual corrections of word boundaries upon which refined segment alignments were created. The study reveals that (i) FL is a widespread process across languages; (ii) FL shows a wide variety of manifestations with respect to the degree and scope of lengthening; (iii) there are several significant interactions between phonological length and positional lengthening. These results lend support to theories assuming a phonological nature of Final Lengthening.

ANNOTATION DECISIONS

STOP ONSET

- Sentential frame context ‘I say X again’: in both languages, target item is followed by adverbial tɔ́na ‘again’

(9) mɛ́ ndɔ la X **tɔ́na.**
/mɛ ^{H}ndɔ la X **tɔ́na**/
SM.1SG PRS say X again
'I say X again.' (Tunen)

ANNOTATION DECISIONS

- Boundary between target and following /t/ based on drop-off of energy on spectrogram

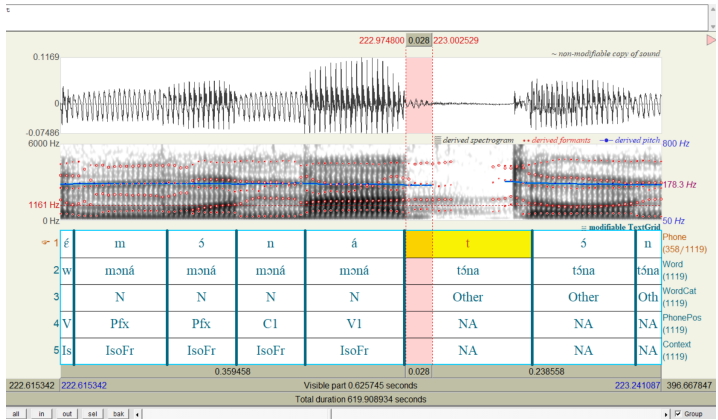
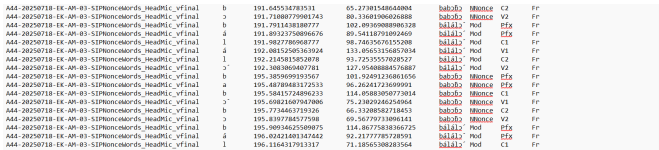


Figure: Annotation of boundary between end of target word and following /t/: highlighted part is not included in duration of vowel (NB: tɔ́na normally not annotated; marked here for expository purposes).

TEXTGRID ANNOTATIONS → ANALYSIS

- TextGrid → .txt file via Praat script:[1]
  - Creates log.txt file with Tier data for all .TextGrid files in the directory
  - .txt file used for statistical analysis (via Excel/R)



```
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      b       191.645534783531        65.2730154864004        babəb   NNonce  C2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      ɔ       191.7108077901743       80.3360190602688        babɔb   NNonce  V2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      b       191.7911438180777       102.0936908806328       bilil   Mod     PFx     Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      á       191.8932375809676       89.5411879102460        bilíl   Mod     PFx     Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      l       191.9827786968777       98.7463567155208        bilíl   Mod     C1      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      á       192.0015250536924       133.0565315685704       bilíl   Mod     V1      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      l       192.2145815852078       93.72535557029527       bilíl   Mod     C2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      ɔ       192.3083060407781       127.9540888476887       bilíl   Mod     V2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      b       195.3859699193567       101.9249123680656       babɔb   NNonce  PFx     Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      a       195.4878948137253       96.26241723699991       babɔb   NNonce  PFx     Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      b       195.5841572489233       114.0588305077014       babɔb   NNonce  C1      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      ɔ       195.6902160794706       75.23029426254964       babɔb   NNonce  V1      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      b       195.7734463719326       66.33208582718453       babɔb   NNonce  C2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      ɔ       195.8397784577598       69.56779731096141       babɔb   NNonce  V2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      b       195.9093462550975       114.8677583836725       bilíl   Mod     PFx     Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      á       196.0242314013474       92.21777785728591       bilíl   Mod     PFx     Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      l       196.1164317933317       71.18565308283564       bilíl   Mod     C1      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      á       196.1876174441454       108.3963353761419       bilíl   Mod     V1      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      l       196.2960813779970068    73.61243671067541       bilíl   Mod     C2      Fr
A44-20250718-EK-AM-03-SIPNonceWords_HeadMic_vFinal      ɔ       196.3696262165013б      91.4088499813488        bilíl   Mod     V2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       17.4005592132454616     105.4666825133452З      nima    N       C1      Iso
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       17.5148258957587б       108.5240006993757З      nimá    N       V1      Iso
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       17.6233498964581З7      89.372706458306О8       nima    N       C2      Iso
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       17.7127226029164643     225.8256779259О424      nimá    N       V2      Iso
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       26.1461229585489б       99.688077048066З3       nima    N       C1      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       26.2442931728048б       89.2664570877З34        nimá    N       V1      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       26.3335906299536З3      95.9614413818285        nima    N       C2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       26.4295210712918б       98.1931027646503З       nimá    N       V2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       28.8068093846108б       106.153058832266З1      nima    N       C1      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       28.9102424434422В2      106.153058832226З6      nimá    N       V1      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       29.0163955022745З5      82.1184060231907        nima    N       C2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       29.0985159062768б4      95.1371753685386        nimá    N       V2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       31.8385829380959б3      144.875843404555В4      nima    N       C1      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       31.9034587816422З       97.6844535934289З      nimá    N       V1      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       m       32.0811421685795        104.715639951173З      nima    N       C2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       á       32.1858588560808В2      101.2000376722О946      nimá    N       V2      Fr
A44-20250718-EK-EB-01-SIPIsolation_HeadMic_vFinal       b       42.2275594960374О5      35.0608808975З38246     bumimá  N       PFx     Iso
```

Ln 2, Col 105 340.354 characters Plain text 100% Windows (CRLF) UTF-16 BE

Figure: Extract from log.txt file generated by Praat script.

[1] With thanks to Connor McCabe (University College Dublin).

RESULTS

RQ: Do Tunen and Nomaandé show lengthening of C_1?

→ Check for length of /p-b/ and /m/ in C_1 vs Pfx, C_2, C_{3+} position, for each language

→ Check for influence of other factors (speaker, word category, discourse context)

BILABIAL STOPS: P OR B?

- Tunen and Nomaandé community orthographies use <b>, but there is no phonological p/b contrast and sources report that the phonetic realisation of plain bilabial stops is usually [p] (Boyd 2015; Philippson 2022; Kerr 2024)
  - ‘/p-b/’ used til now in this presentation as placeholder
  - Voicing to be checked in Praat

BILABIAL STOPS: P OR B?

- VOT continuum
  - VOT < 0: voiced consonant (**b**)
  - VOT = 0: voiceless unaspirated consonant (**p**)
  - VOT > 0: voiceless aspirated consonant (**p^h**)

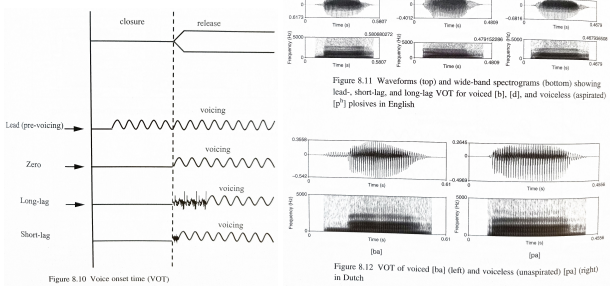


Figure 8.10 Voice onset time (VOT)

Figure 8.11 Waveforms (top) and wide-band spectrograms (bottom) showing lead-, short-lag, and long-lag VOT for voiced [b], [d], and voiceless (aspirated) [p^h] plosives in English

Figure 8.12 VOT of voiced [ba] (left) and voiceless (unaspirated) [pa] (right) in Dutch

Figure: Illustration of VOT difference (Wayland 2019).

BILABIAL STOPS: P OR B?

- Tunen and Nomaandé have /**p**/ ($\rightarrow$ [b] / m_): lack of voicing bar before stop release; VOT $\approx$ 0

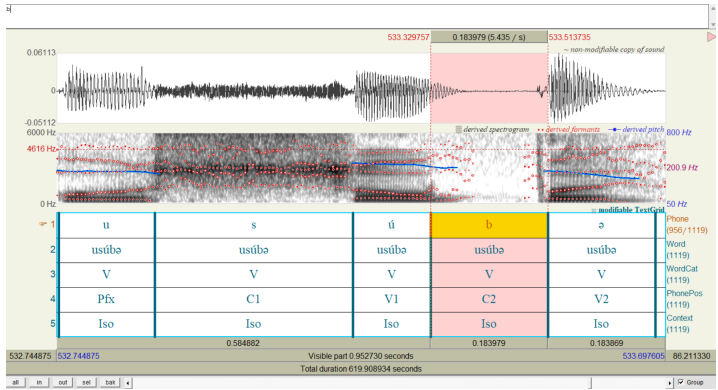


Figure: Example of voiceless unaspirated realisation of word-medial /p/ in Tunen verb usúbə 'to drum'.

BILABIAL STOPS: P OR B?

- Tunen and Nomaandé have /**p**/ ($\rightarrow$ [b] / m_): lack of voicing bar before stop release; VOT $\approx$ 0
- Stop length therefore cannot be checked when stop /p/ is in initial position as there is no voicing bar that can be identified on the spectrogram and used to measure length of pre-closure part of stop

$\rightarrow$ Exclude /p/ from main analysis due to non-comparability with /b/ in languages of Idiatov and Van de Velde (2016)'s sample

ANNOTATING /M/

- /m/ selected due to ease of identification from spectrogram

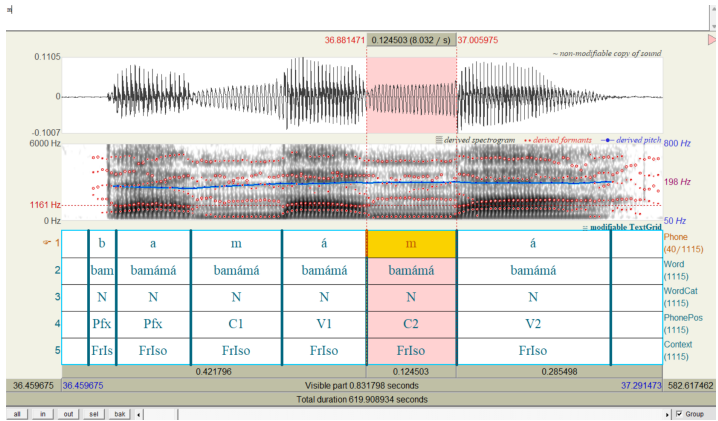


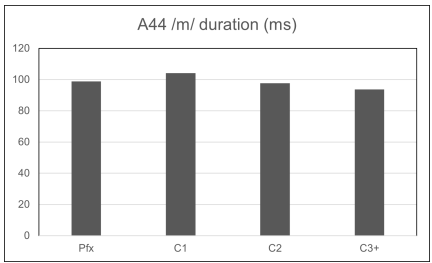
Figure: Example of annotation of /m/ in Praat.

PRELIMINARY RESULTS

- Annotation ongoing...
  - 3117 annotated phones for Tunen (4 speakers), of which 574 /m/, 285 /p/
  - 366 annotated phones for Nomaandé (1 speaker), of which 65 /m/, 33 /p/

PRELIMINARY RESULTS

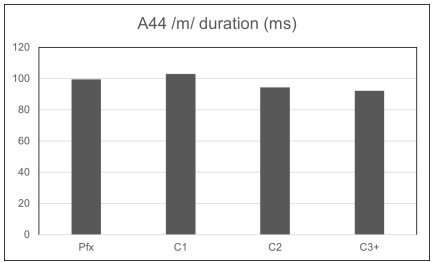
- Mean duration of /m/ in Tunen (across speakers):



Position	Duration (ms)	Count
Pfx	98.90	186
C1	104.20	181
C2	97.64	175
C3+	93.71	32
		574

PRELIMINARY RESULTS

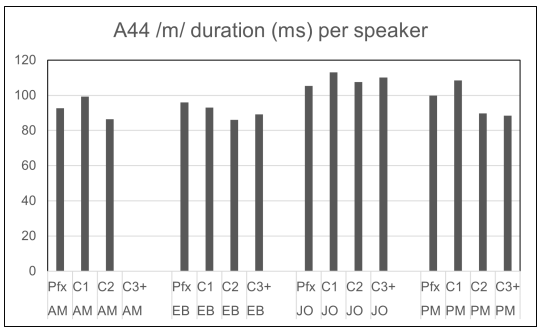
- Mean duration of /m/ in Tunen (across speakers, excluding isolation context (Iso)):



Position	Duration (ms)	Count
Pfx	99.54	137
C1	103.01	133
C2	94.28	130
C3+	92.24	24
		424

PRELIMINARY RESULTS

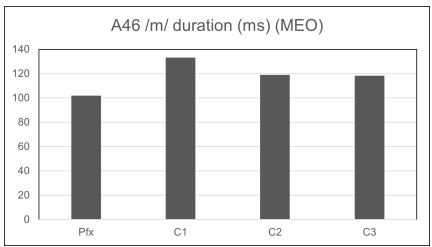
- Split per speaker (excluding isolation context (Iso)):



Consultant	Position	Duration (ms)	Count
AM	Pfx	92.75	15
AM	C1	99.31	19
AM	C2	86.47	20
AM	C3+	-	0
EB	Pfx	96.01	51
EB	C1	93.10	50
EB	C2	86.11	49
EB	C3+	89.21	9
JO	Pfx	105.40	47
JO	C1	113.15	46
JO	C2	107.63	46
JO	C3+	110.06	8
PM	Pfx	99.82	24
PM	C1	108.48	18
PM	C2	89.72	14
PM	C3+	88.42	6
			422

PRELIMINARY RESULTS

- Preliminary analysis for Nomaandé (consultant MEO):



Position	Duration (ms)	Count
Pfx	102.02	19
C1	133.13	20
C2	119.03	22
C3	118.21	4
		65

EXAMPLE: /M/ LENGTH IN REAL WORDS

- No visible length difference of /m/ in real words

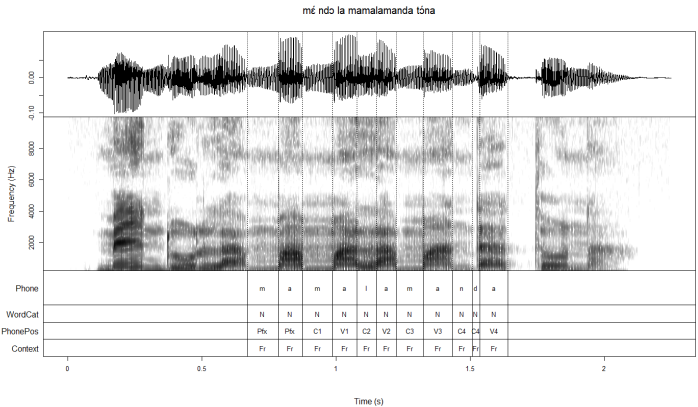
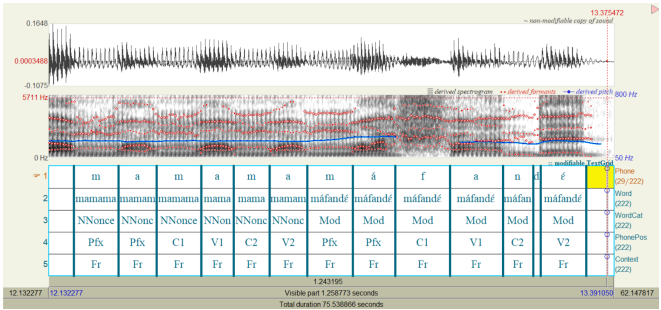


Figure: Example of comparable /m/ length in prefixal, C_1, and C_3 positions (Tunen; Fr context). Figure made in R using Praatpicture (Rasmus Puggaard-Rode, University of Oxford).

EXAMPLE: /M/ LENGTH IN NONCE WORDS

- No visible length difference of /m/ or /a/ in nonce words



1		m	a	m	a	m	a	m	á	f	a	n	d	é
2		mamama	mamam	mamama	mama	mama	mamam	máfandé	máfandé	máfandé	máfandé	máfan		máfandé
3		NNonce	NNonc	NNonce	NNon	NNonc	NNonc	Mod	Mod	Mod	Mod	Mod		Mod
4		Pfx	Pfx	C1	V1	C2	V2	Pfx	Pfx	C1	V1	C2		V2
5		Fr	Fr	Fr	Fr	Fr	Fr	Fr	Fr	Fr	Fr	Fr		Fr

Figure: TextGrid annotation for Tunen nonce word mamama in context of numeral máfandɛ ‘two’.

EXAMPLE: /M/ LENGTH IN CONTRASTIVE CONTEXT

- No visible evidence of C-emphasis in corrective focus context

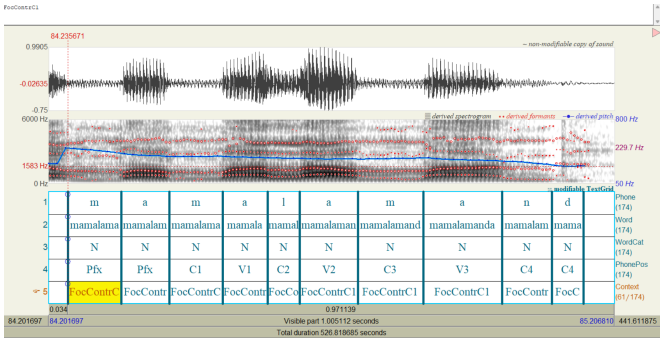


Figure: TextGrid annotation for Tunen word mamalamanda ‘thunders’ in context of correction of mispronunciation of C1 by EK.

WHAT ABOUT VOWEL LENGTH?

- Vowel length annotations more complicated, e.g. vowels in final position are consistently lengthened
- Preliminary analysis of non-final vowels: no evidence that V1 is lengthened

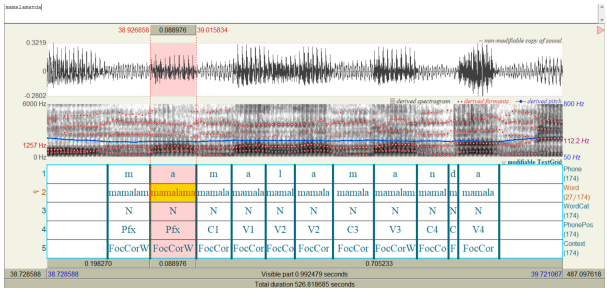


Figure: Example length of /a/ in Pfx, V1, V2, V3, and V4 positions in Tunen word mamalamanda ‘thunder (pl.)’ in corrective focus context on the word level.

PRELIMINARY RESULTS: SUMMARY

! Data annotation & statistical analysis still ongoing

- Preliminary analysis does not suggest lengthening of /m/ in C1 position, or only very weakly

- Stats to do
  - Analysis to check for confounding variables

  → How to control for variation in speech rate?

? What is the baseline to compare against?

DISCUSSION AND CONCLUSION

GHENT
UNIVERSITY

SUMMARY

MAXIMALITY CONSTRAINTS

- Maximality restrictions = phonological restrictions on size of verb stem (verb root + suffixes)
- Cameroonian Bantu/Bantoid generally considered to have shorter verb stems than other Bantu
- Variation within Cameroonian Bantu/Bantoid in strength of maximality restrictions
- Preliminary results for Mbam languages show weaker evidence for maximality constraints than in surrounding Bantoid languages
  - e.g. can stack DIM suffix

SUMMARY

STEM-INITIAL PROMINENCE

- Stem-initial prominence (SIP) = phonological restrictions prominence on beginning of verb stem
- SIP may drive maximality constraints, under diachronic scenario synthetic verb $\rightarrow$ breakdown, which in turn leads to var. changes in verbal morphosyntax (Hyman 2004, 2017, 2018)
- SIP can be seen through phoneme inventories and may be visible through phonetic effects e.g. lenghtening of C1 and V1
- Idiatov and Van de Velde (2016) report C1 lengthening for 7/9 Cameroonian Bantu/Bantoid languages
- Replication study for neighbouring languages Tunen and Nomaandé did not show convincing evidence for C1 lengthening

DISCUSSION

AREAL DISTRIBUTION OF SIP

- NB: Lack of phonetic correlates of SIP in Tunen and Nomaandé does not necessarily mean no SIP (e.g. visible in phoneme inventory)

- Strength of the lengthening effect in Idiatov and Van de Velde (2016) study reported to decrease in the south of the A70 group, i.e., to be less visible in the Fang varieties of Gabon than the A70 languages of central Cameroon
- Lack of lengthening found for Bapuku and Orungu suggested by Idiatov and Van de Velde (2016) to be linked to broader phonological and morphosyntactic differences in the grammar of these languages

- Lovegren (2012) on Mundabli (Bantoid, Cameroon): evidence for SIP from phoneme inventory, but no phonetic correlates w.r.t. length $\rightarrow$ SIP is “covert”

DISCUSSION

AREAL DISTRIBUTION OF SIP

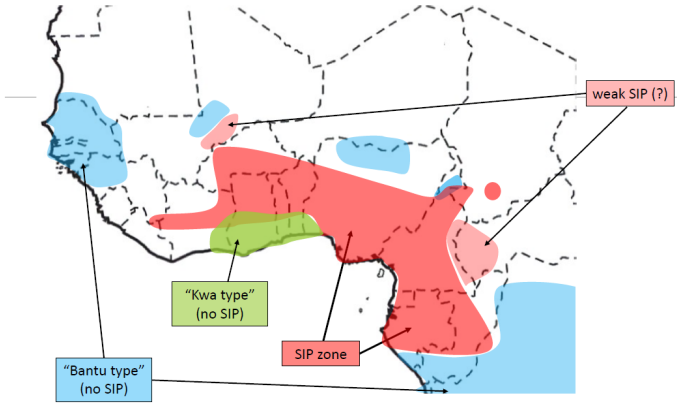


Figure: Preliminary map of strength of SIP by Lionnet (Lionnet 2017:46)

DISCUSSION

- Mbam Bantu show clear morphosyntactic separation between pre-stem material and verb stem (split predication), i.e. analyticity wrt integration of pre-stem material (Kerr to appear)

(10) Tɔ ŋa cáŋa esú **súéte**.
SM.1PL PST3 first PRN.1PL leave
`We left first.' (Nomaandé; Taylor, 1999, §1.0, adapted)

- But Mbam languages found here to be less analytic in the sense of weak maximality constraints/SIP effects

⇒ independence of pre-root analyticity and post-root analyticity

CONCLUSION

- Cameroonian Bantoid verb stems vary in multiple dimensions
  - Degree of separation of verb stem from pre-stem material (Kerr to appear)
  - Degree of suffixal expansion of verb form (maximality constraints)
  - Degree of stem-initial prominence effects (stem-initial prominence)

- Areal distribution is heteregeneous
  - ! Moreover, not a clear North/South pattern
  - Variation within the ‘SIP zone’ identified for Cameroon in Lionnet (2017) preliminary overview

THANK YOU!

Acknowledgements:

All Tunen and Nomaandé consultants in Ndikiniméki/Bokito.

Dmitry Idiatov for sharing stimuli; Matthew Faytak, Dmitry Idiatov, Connor McCabe for advice on annotations; Connor McCabe for creating custom Praat script. CNRS-LLACAN for hosting research stay. All mistakes/annotation errors my own.

Financial support: Research Foundation - Flanders (FWO) project FWO-MBAM (1274225N; PI Elisabeth J. Kerr, Ghent University) long stay abroad grant; Foundation for Endangered Languages (FEL); Endangered Language Fund (ELF).

REFERENCES I

Boyd, V. L. (2015). The phonological systems of the Mbam languages of Cameroon with a focus on vowels and vowel harmony. PhD thesis, Leiden University.

Good, J. (2012). How to become a "Kwa" noun. Morphology, 22:293–335.

Good, J. (2017). Niger-Congo languages. In Hickey, R., editor, The Cambridge Handbook of Areal Linguistics, pages 471–499. Cambridge University Press, Cambridge.

Grimm, N. (2025). Towards Reconstructing Grammatical Tone in the Northwestern Bantu Verb Phrase. Diachronica, 42(3):32456.

Guérois, R. (2015). A grammar of Cuwabo (Mozambique, Bantu P34). PhD thesis, Université Lumière Lyon 2.

Güldemann, T. (2022). Predicate and argument indexing in early Bantu. In Bostoen, K., de Schryver, G.-M., Guérois, R., and Pacchiarotti, S., editors, On reconstructing Proto- Bantu grammar, page 389424. Berlin: Language Science Press.

Hyman, L. M. (2004). How to become a Kwa verb. Journal of West African Languages, 30(2):69–88.

Hyman, L. M. (2017). Multiple argument marking in Bantoid: From syntheticity to analyticity. In Malchukov, A. and Bisang, W., editors, Unity and diversity in grammaticalization scenarios, pages 67–95. Language Science Press, Berlin.

Hyman, L. M. (2018). Common Bantoid verb extensions. In Watters, J. R., editor, East Benue-Congo: Nouns, pronouns, and verbs, pages 173–198. Language Science Press.

REFERENCES II

Hyman, L. M., Rolle, N., Sande, H., Clem, E., Jenks, P., Lionnet, F., Merrill, J., and Baier, N. (2019). Niger-Congo linguistic features and typology. In Wolff, E., editor, The Cambridge Handbook of African Linguistics, pages 191–245. Cambridge University Press.

Idiatov, D. and Van de Velde, M. L. O. (2016). Stem-initial accent and C-emphasis prosody in north-western Bantu.

Kerr, E. J. (2024). Tunen syntax and information structure. PhD thesis, Leiden University.

Kerr, E. J. (to appear). Auxiliaries and analyticity in Northwest Bantu. Nordic Journal of African Studies.

Lionnet, F. (2017). Stem-initial prominence in West and Central Africa: Niger-Congo, areal, or both? Paper presented at the 48th Annual Conference on African Linguistics (ACAL 48), Indiana University, Bloomington 30 March - 2 April 2017.

Lionnet, F. and Hyman, L. (2018). Current issues in African phonology. In Güldemann, T., editor, The languages and linguistics of Africa, pages 602–708. de Gruyter, Germany.

Lovegren, J. (2012). Stem-initial prominence in Mungbam. In Marlo, M. R., Adams, N. B., Green, C., Morrison, M., and Purvis, T. M., editors, Selected Proceedings of The 42nd Annual Conference on African Linguistics. Cascadilla Proceedings Project.

Meeussen, A. E. (1967). Bantu grammatical reconstructions. Africana linguistica, 3(1):79–121.

Nurse, D. (2008). Tense and aspect in Bantu. Oxford: Oxford University Press.

REFERENCES III

Paschen, L., Fuchs, S., and Seifart, F. (2022). Final lengthening and vowel length in 25 languages. Journal of Phonetics, 94:101179.

Philippson, G. (2022). Double reflexes in northwestern Bantu and their implications for the Proto-Bantu consonant system. In Bostoen, K., Pacchiarotti, S., Guérois, R., and de Schryver, G.-M., editors, On reconstructing Proto-Bantu grammar, pages 3–58. Berlin: Language Science Press.

Seifart, F., Strunk, J., and Danielsen, Swintha, e. a. (2021). The extent and degree of utterance-final word lengthening in spontaneous speech from 10 languages. Linguistics Vanguard, 7(1):20190063.

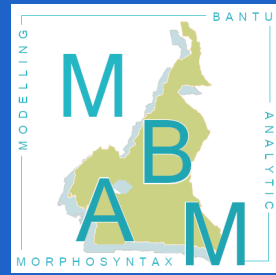
Zahran, A., Devos, M., and Guérois, R. (to appear). The development of negative 'have' auxiliaries in the Bantu languages of the Middle and Lower Zambezi. Nordic Journal of African Studies.

FACULTY OF ARTS
AND PHILOSOPHY

Elisabeth J. Kerr

FWO Junior Postdoctoral Fellow
Department of Languages and Cultures
Ghent University (Belgium)

: Elisabeth.Kerr@UGent.be
: https://research.flw.ugent.be/en/elisabeth.kerr



GHENT
UNIVERSITY

BantUGent

fwo

Langage, Langues et
Llacan
Cultures d'Afrique