Lobi stem alternations: Two morphophonological puzzles

Hannah Sande[1]
UC Berkeley

1 Introduction

In Lobi (iso:lob; Gur/Mabia; Burkina Faso, Ghana, and Côte d'Ivoire), most nouns and adjectives have two allomorphs, a *long form* and *short form*, depending on their position in the noun phrase.

- This talk considers two puzzles that arise in describing and analyzing Lobi stem alternations.

Puzzle #1: What is the nature of the relationships between the long and short forms of nouns and adjectives? Is one predictably phonologically derived from the other, or must they both be lexically listed?

- I show that the long forms are not predictably derivable from the short forms.
- However, despite apparently messy and unpredictable surface facts, the short forms are predictably derived from the long forms via truncation.
  - This characterization of the pattern differs from previous treatments of Gur/Mabia stem alternations, which either assume that the short form is basic (cf. Miehe and Tham 2007) or treat the two forms as suppletive (Sande To Appeara).
- I argue that stems shorten to meet particular output requirements in certain positions.
- The data and analysis have implications for our understanding of how truncation is formally derived:
  - Truncation in Lobi can be blocked if it would lead to violations of other phonotactic constraints, which suggests that truncation is best analyzed as a violable constraint on the prosodic shape of outputs.
    - Along the lines of Kurisu (2001)
    - Contra, e.g., McCarthy and Prince (1996, 1999)

Puzzle #2: What is the correct generalization about where long vs short forms appear? Is it prosodic or morphosyntactically determined? The realization of some morpheme?

- I show that the distribution of long versus short forms is not prosodically predictable nor is it the realization of some morpheme, but it may be morphosyntactically predictable.
- How to formalize the morphosyntactic generalization remains a puzzle.

[1]Endless thanks to Sansan Claude Hien, who has devoted so much time and knowledge over the past five years to the Lobi Documentation Project. Thanks also to Kambou Sié Moïse, Dah Frank, Dah Doudou, and Kambiré Bebe Hien for their contributions. Thanks to the students of the Berkeley undergraduate field methods class in Spring 2022, the year-long grad field methods class in 2022-2023, and the Fall 2025 undergraduate morphology class for learning about this language alongside me. Thanks to Dylan Worden and audiences at AMP 2026 and the Fall 2026 "What African languages can teach us about phonology" workshop for helping me to better understand Lobi plural morphology, which is related to the facts presented here. And thanks to Florian Linonet and Mirella Blum for their amazing talks at ACAL 57 that inspired me to re-look at this complex data in a new way. The analysis presented here is my own, and any errors are mine. AI was not used in the data collection, analysis, or write-up for this talk.

2 A brief background on truncation and subtraction

While the canonical realization of some morpheme is the addition of one or more segments, there are also morphological processes that remove or delete segments.

- These fall into two categories (Weeda 1992; Kurisu 2001; Inkelas 2014):
  i. Templatic truncation[2]
  ii. Subtraction
- Subtraction involves removing material of a consistent size (e.g., remove the final segment).
- Templatic truncation involves a consistent amount of structure left behind after deletion (e.g., remove everything but a foot) (Kurisu 2001, p. 7).
- Both truncation and subtraction seem to target prosodic constituents like a prosodic word, foot, syllable (light, heavy, or unrestricted), or mora (McCarthy and Prince 1996, 1999).

Truncation and subtraction have been analyzed in a variety of ways:

- As a morphological process in a purely process-based morphological framework (Anderson 1992).
- As templatic affixation in the Theory of Prosodic Morphology (McCarthy and Prince 1996, 1999)
  - Templates are affixes that are made of units of the prosodic hierarchy.
  - The Satisfaction Condition: All elements in a template are obligatorily satisfied.
  - For other examples of using prosodic templatic affixes to derive truncation or subtraction (Trommer and Zimmermann 2014; Zimmermann 2017).
- Via universal constraints like Affix = Syllable, Stem = Foot in Generalized Template Theory (Urbanczyk 1996, 2006; Downing 2006).
- With the constraint ranking RealizeMorph ≫ Max (Kurisu 2001).
  - Kurisu (2001) proposes that both subtraction and truncation involve the ranking RealizeMorph ≫ Max.
  - They differ in that truncation also involves "a set of prosodic constraints restricting the phonological size of the entire output" that outrank Max (p. 25).

I will argue that the stem alternations in Lobi involve truncation, and I propose an analysis along the lines of Kurisu (2001), where phonotactic constraints on the size of some constituent of the output outrank Max.

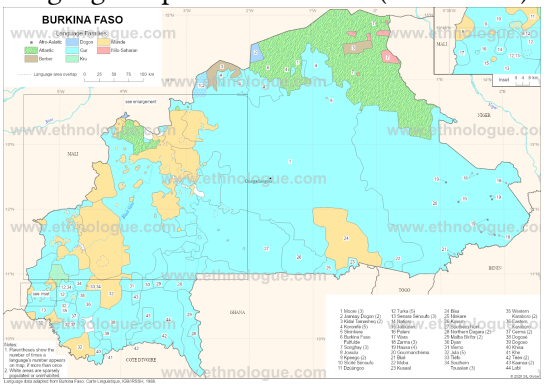
- For Lobi, we don't need RealizeMorph because, as I will argue, the relevant process is not the realization of any morpheme.
- While McCarthy and Prince (1996, 1999) claim that templates are always met, I will show that their satisfaction can be blocked just in the case that satisfying the template would violate a higher-ranked constriant.
  - This suggests that a constraint-based analysis of truncation is preferable to an affixation analysis.

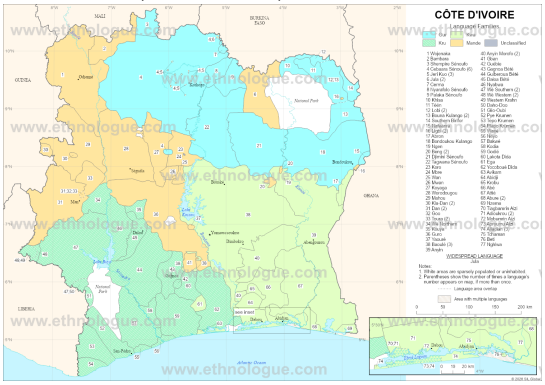
[2]Kurisu (2001, p. 79) claimed that templatic truncation, while found in language games and nickname formation, is never grammaticalized as the realization of some morpheme. However, there have since been several documented cases of truncation as the realization of a morpheme, for example, in Amahuaca Clem (2019) and Rarámuri (Caballero 2008).

3 Background on the Lobi language

Lobi is a Gur (also called Mabia) language spoken in Burkina Faso, Côte d'Ivoire, and Ghana.

(1) Language map of Burkina Faso (Lobi is # 29) and Côte d'Ivoire (Lobi is # 12)





- Ethnologue (Eberhard et al. 2026) lists Lobi as having a total of 700,600 speakers.
- Miehe and Tham (2007) estimate that there are 300,000 people in Burkina Faso and Côte d'Ivoire who speak Lobi, with a rather small population of Lobi speakers in Ghana.
- Many Lobi people in Burkina Faso and Côte d'Ivoire have been displaced from their homes during recent national crises, resulting in groups of Lobi people moving south and settling in traditionally non-Lobi villages throughout southern Côte d'Ivoire.

The data presented here was collected as part of an ongoing Lobi documentation project that began in a field methods course at UC Berkeley in 2022.

- Speakers: Five total Lobi speakers' voices are represented in the corpus, though most of the data comes from a single speaker, Sansan Claude Hien.
  - Hien is from Agnibilékrou, Côte d'Ivoire and has lived in the US since 2015. He is currently pursuing a Masters degree in Public Policy at UC Berkeley.
  - The other Lobi speakers represented in the corpus are Dah Doudou and Dah Frank from Gnagbodougnoa, Côte d'Ivoire, and Moïse Sié Kambou, and Kambiré Bebe Hien from Abidjan, Côte d'Ivoire.

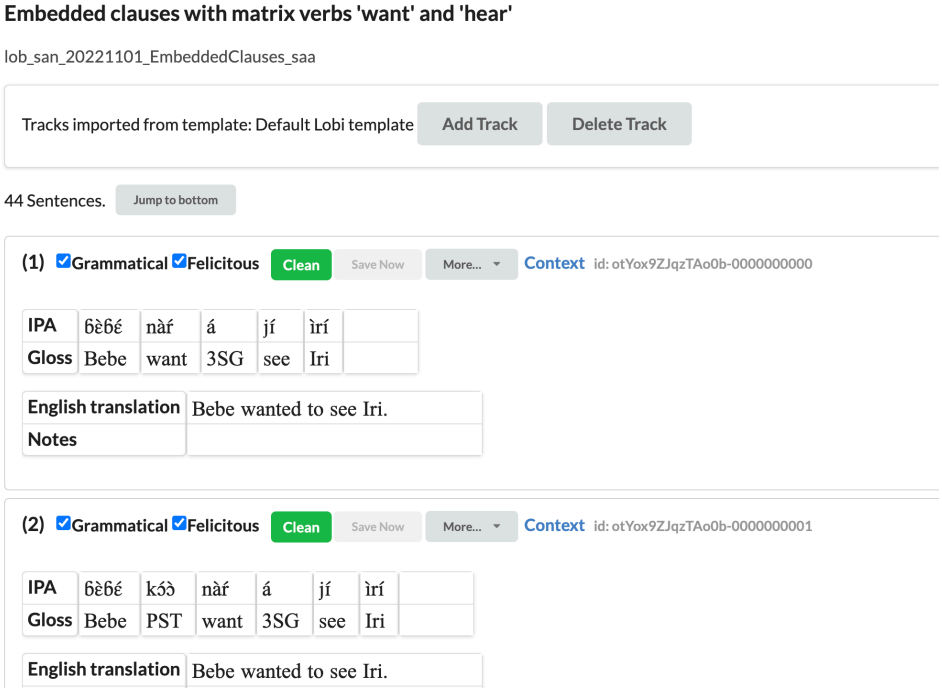
(2) Photos of the Spring 2023 Berkeley Field Methods class with Sansan Claude Hien and Hannah Sande working with Dah Doudou in Gnagbodougnoa, Côte d'Ivoire





- Corpus: The project has resulted in a corpus of roughly 9,000 transcribed, glossed, and translated utterances, which inform the generalizations presented here.
  - The corpus is maintained using Twisted Tongues (Sande and Ewert 2026), a database management tool that makes it easy to add, organize, search, and export language data:

(3) A screenshot of a passage in Twisted Tongues



- Dataset: In particular, I draw generalizations from a list of nearly 400 nouns for which we have data on the singular, plural, long, and short forms.
  - All items in the list were recorded with Lobi speaker and collaborator Sansan Claude Hien, elicited by me and Berkeley undergraduate student Dylan Worden during 2026.
  - Additionally, a substantial portion of the list was checked with a second Lobi speaker, Moïse Sié Kambou, in Fall 2026.

(4) Hannah Sande with Moïse Sié Kambou in Abidjan, Côte d'Ivoire

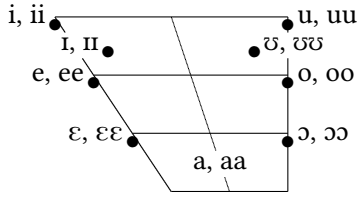


This talk focuses on the long and short stem alternations among nouns and adjectives in Lobi, though I first give brief background on its vowel system, syllable structure, tone, and the morphosyntax of noun phrases, which are relevant for understanding the stem alternation facts.

3.1 Lobi vowels

Lobi has 9 contrastive vowel qualities, all of which also show a length contrast.

(5) Vowel inventory



(6) Vowel features

	i	ɪ	e	ɛ	a	u	ʊ	o	ɔ
high	+	+	-	-	-	+	+	-	-
low	-	-	-	-	+	-	-	-	-
round	-	-	-	-	-	+	+	+	+
ATR	+	-	+	-	-	+	-	+	-

- There are no contrastive nasal vowels, but vowels are predictably nasalized before a nasal coda.[3]
- Vowels and sonorant consonants nasalize in suffixes or clitics after following a nasal coda consonant.
- Long vowels can be underlying or derived.
- Within a morpheme, possible diphthongs include only a high vowel followed by a mid vowel with the same ATR and rounding values. These only surface in word-final syllables closed by a coronal coda consonant:
  - [bɪɛl] 'one'
  - [ɲʊ̃ʊ̃n] 'water'
  - [falcɪɛr] 'broom'
  - [gʊ̃ngʊ̃ʊ̃n] 'hill/mountain'

3.2 Syllable structure

Possible surface syllable structures in Lobi include those listed in 7.

(7) Possible surface syllable structures

a.	CV	na	'cow'
b.	CVV	síí	'snake'
c.	CVL	bʊkár	'stone'
d.	CVɓ	$k^huɓ$	'bone'
e.	CLV	bló	'white'
f.	CVVL	ɲʊ̃ʊ̃n	'water'
g.	N	ń-ní	'INAL-mother'

[3] In previous work, Labouret (1958) and Miehe and Winkelmann (2007) list nasal vowels in underlying forms. Becuwe (1982); Kambou (2003); Vaillant (1967) and Lamothe (1964, 1966) consider nasal vowels to be derived, as I do here.

- Possible codas: Oral and nasal sonorants and the implosive /ɓ/ (not obstruents or glottalized sonorants).
  - I've separated CVɓ from other CVL syllables because CVL syllabes can bear two tones but CVɓ syllables can only bear one.
- The second consonant in CLV syllables can only be an alveolar sonorant, [r,l,n].
- The first consonant in CLV syllables can be a fricative or (oral or nasal) stop.
- Syllabic nasals are rare and are always functional morphemes.
- Before a word-final coda [n] (but not other nasals), vowels systematically lengthen and nasalize. The final [n] is only variable pronounced: /bín/, [bı́ı̃ı̃(n)] 'dog'.
  - This can result in alternations between short and long vowels, as in [kontı́ı̃ın] 'elder', [kon.tí.n-ā́] *elder-*PL, 'elders'.

3.3 Tone

There are two contrastive tone heights, L and H.

- Ex: tʰii 'medicine' vs tʰíí 'soil'
- Contrastive tone melodies on one- or two-syllable roots include H, L, and LH.
  - HL is robustly attested in derived contexts but not underlyingly.
- Here, I write H tones with an acute accent and leave L-tones unmarked unless in a contour.

The tone-bearing unit is the mora.

- CV and CVɓ syllables can only host a single tone height, while CVV, CVL, and CVVL syllables can host two tones heights.
- This suggests that the tone-bearing unit is the mora.
- Vowels and sonorant codas can also bear tone: [wɛ̀Ĺ] 'pepper'; [mı̂ı̂n] '1SG.POSS'.
- There is non-automatic downstep at certain junctures (Sande and Hien To Appear), and certain grammatical morphemes introduce floating H tones that surface on the first TBU of the next morpheme (Sande To Appearb).

3.4 Nominal morphosyntax

(8) Nominal morphology template
POSS *noun*-PL=DEM=DEF

(9) Noun phrase word order
Possessor Noun Adj Num Quant(=)Dem=Def

- Number can be marked morphologically or by a separate generalized plural word /sʊ̀nɔ/.
  - Abstract nouns cannot be pluralized.
  - Nouns that have a collective meaning by default (things that typically come in pairs/groups), or that are unique, cannot take a morphological plural but can surface with the generalized plural.[4]

[4]Speaker Kambou used many more instances of the generalized plural word than Speaker Hien. Miehe and Tham (2007) note based on comparing work on Lobi over time that the generalized plural word seems to be becoming more common, so Hien's speech perhaps represents a more conservative variety.

  - Nouns that are singular in their default form can take a morphological plural.[5]

(10) A sample of Lobi morphological plurals

	Singular	Plural	Meaning
a.	tibil/tibli	tɪbla	'person'
b.	mãnɟ́ĩ́ĩ́n	mãnɟɛ́nã́	'papaya'
c.	ɓũndé	ɓũndɛ́	'rat'
d.	kʰuɓ	kʰʊɓa	'bone'
e.	bʊ	bɔ	'goat'
f.	kór	kórá	'hole'
g.	nɔ́ɔ́	nɛ́ɛ́	'leg'
h.	nũntodáàr	nũntodára	'earring'
i.	gɔ̃ngʊ̃ɔ̃n	gɔ̃ngɔnã	'hill/mountain'

- Morphological plural formation appears messy, though Sande and Worden (2026) show that plural realization is phonologically predictable from a single underlying form: a floating moraless bundle of /a/-like features.

Most nouns and adjectives have two distinct morphological forms, whose distribution is morphosyntactically predictable.

(11) Example long and **short** forms of 'house' in context (20220328_ngn, 20230725_hls)

a. cʊ̀ɔr
house
'house'

b. cʊ̀ɔr=rá
house=DEF
'the house'

c. **cɔ** bɪɛl
house one
'one house'

d. mĩ́ĩ̀n ↓**cɔ́**
1SG.POSS house
'my house'

4 Long and short forms

Long/short allomorphy applies to both singular and plural nouns, and to adjectives.

- This section focuses on the long and short *forms* and whether one is predictable from the other.
- I start with singular nouns and discuss their alternations in form.
  - I don't have time to talk about plural nouns today, but feel free to ask.
- I show that the alternations are the same among adjectives.
- I provide a constraint-based analysis to account for the alternations, which I argue involve truncation.

[5]In most Gur/Mabia languages, morphological number is marked through suffixing singular and plural noun class markers (Miehe and Winkelmann 2007; Miehe et al. 2012). The noun class system in Lobi, though, has been (almost) entirely lost (cf. Miehe and Tham (2007).

4.1 Singular nouns

- The long and short forms of several nouns are given in (12).

(12) Short and long forms of nouns

	Short form	Long form	
a.	kũ	kũũn	'man'
b.	jí	jír	'eye'
c.	bẽ	bẽm	'deer'
d.	cɔ	cʊɔr	'house'
e.	jɛ	jɪɛr	'year'
f.	ɲɔ̃	ɲʊ̃ɔ̃n	'water'
g.	ɲɔ̃	ɲɔ̃ɔ̃n	'arm'
h.	pú	púúr	'pigeon'
i.	bɛ̃́	bɛ̃́ɛ̃́n	'dog'
j.	lɔ́	lɔ́ɔ́	'farm'
k.	nɔ́	nɔ́ɔ́	'leg'
l.	mãnɟɛ̃́	mãnɟɛ̃́ɛ̃́n	'papaya'
m.	lõmbi	lõmbir	'bird'
n.	bʊ	bʊ	'goat'
o.	pàá	pàá	'leaf'
p.	wɛ̀Í	wɛ̀Í	'pepper'
q.	k^huɓ	k^huɓ	'bone'
r.	kókó	kókó	'monkey'

Types of alternations include:

- short vs. long vowel (12a,g,h,i,j,k,l)
- no coda vs. final consonant (12a,b,c,d,e,f,g,h,i,l,m)
- diphthong vs. singleton (12d,e,f,)
- vowel quality alternations (12i,l)
- some combination of the above (12a,d,e,f,i,l)
- no change (12n,o,p,q,r)

- There is no systematic addition of a particular segment or segments from short to long form.
  - There are minimal pairs of short forms that have different long forms, as in (12f,g).
  - It's also not the case that all long forms share the same output structure.
    - They can be CV (12n), CVV (12j,k), CVC (12b,c), CVVC (12a,d), or longer (12l,m).
    - Any valid syllable structure and word shape in the language is attested among long forms.
- And while there is regularly simplification of syllable structure from the long to short form, on first glance these also do not seem to be predictable or uniform.
  - It is not the case that the same kind or amount of material is always lost between long and short forms.
    - Sometimes one segment, sometimes more.
    - Sometimes a consonant, sometimes a vowel, sometimes both.
  - Some long vowels and diphthongs in the long form correspond with short vowels in the short form (12a,d-l) while others remain long (12o).
  - Some final consonants are present in the long form but not the short form (12a-i,l,m), while others are present in both forms (12p,q).
  - **However,** the short forms have something in common: Their final syllables are light (CV).
    - The exceptions to this (12p,q,r) are principled and due to regular phonological alternations and phonotactic restrictions in the language.

I will show that if the long forms are taken to be basic, short forms are predictably derivable through regular phonological constraints needed elsewhere in the language, plus a context-specific restriction that final syllables be light.

- In general, short forms involve truncation of the final syllable to CV.
- Nouns that fail to alternate do so for the following reasons:

  - They are already CV- or CVɓ–final (12n,q,r)
  - Their final syllable is bimoraic, CVV or CVC, and bears two tones (1o,p)

- The vowel quality alternations between long and short forms are restricted to nasal contexts and are predictable given regular phonotactic restrictions in the language (12i).

4.2 Adjectives show the same patterns

The same form alternations we see among nouns are also found among adjectives.

- Many adjectives, like nouns, have distinct short and long forms (13).
- The same kinds of alternations that we saw among nouns are also present among adjectival long and short forms:
  - vowel length alternations (13a-c)
  - diphthong/monophthong alternations (13d,e)
  - presence of final C (13b-c,e)

(13) **Example long and short forms of adjectives**

	Short form	Long form	Gloss
a.	pʰa	pʰaa	‘new’
b.	gbā́	gbā́ā́n	‘tall’
c.	tī́	tī́ī́n	‘old’
d.	bɔ́	bɔɔ́	‘good’
e.	sɛ	sɪɛr	‘red’
f.	brí	brí	‘black’
g.	bló	bló	‘white’
h.	dɛ̀ɛ́	dɛ̀ɛ́	‘ugly’
i.	pɔ̀Ĺ	pɔ̀Ĺ	‘clean’

- Like nouns, not all adjectives alternate (13f-i).
  - The ones that don’t are either light syllables or CVV/CVC syllabes with a contour tone, as in (18, 19).
- We’ll see in §5 that a single statement about the distribution of long and short forms applies to both nouns and adjectives.
- Given that the generalizations about forms and distributions are the same, a single analysis of long/short forms for nouns and adjectives is possible.

4.3 Short forms are derived from long forms

I argue that, despite the variety of alternations seen in the data, *short forms are predictably derivable from long forms.*

- **Evidence:**
  i. The long form is not predictable from the short form.
     * For example, ‘water’ and ‘arm’ are identical in their short forms, [ɲɔ̃] (12f,g), but the corresponding long forms differ, [ɲɔ̃ɔ̃n] vs. [ɲɔ̃ɔ̃n].
  ii. For nouns that have long/short alternations, all short forms are similar in that they end in a CV syllable.

iii. There are no contrastive nasal vowels in the language–the only place where nasal vowels surface without an adjacent nasal consonant is in short forms.
   * Ex: [kũ] (short) and [kũũn] (long) from /kun/ 'man'

iv. Speakers have judgments about the "correct" short form of a word despite not having heard it before.
   * Ex: Speaker Hien provided a short form for the word 'barren women', [kɔ̃mbɔ́], saying that's what he would say but he can't remember having heard it before.
   * I ran an informal wug test with Hien:

(14) Nonce word results

	Long form (provided)	Short form (produced)	Predicted?
a.	tʰuɓ	tʰuɓ	✓
b.	bil	bi	✓
c.	wẽẽn	wẽ	✓
d.	fɔ̃m	fɔ̃	✓
e.	mor	mo	✓
f.	tʰùú	tʰùú	✓

   * Hien had very clear judgments about these nonce forms, and even made comments like, for example (14f): '[tʰùú], it's not going to be short'.

- **Analytical basics:**
  - I analyze the short forms as derived from the long forms via constraint interaction, and crucially via phonotactic markedness constraints on outputs.
  - I take the fact that all short forms of alternating nouns (and many non-alternating ones (12n,p)) end in a light syllable as evidence for a constraint prohibiting heavy syllables in certain positions.
    * *FinalHeavyσ: Assign a violation for a heavy final syllable in the relevant position.
      · The relevant position, which would need to be baked into this constraint, is discussed in §5.[6]
  - The final light syllable requirement can be overridden if a heavy syllable is needed to bear a two tones (as in 12o), and input tones must be pronounced.
    * *2Tones-1μ: Assign a violation for each mora that is associated to two or more tones.
    * Max-Tone: Assign a violation for each tone in the input that is not present in the output.

- **The OT analysis:**
  - Using the two phonotactic Markedness constraints above, plus Faithfulness constraints Max and Dep, ranked as below, we can derive the correct outputs.
  - *2Tones-1μ, Max-Tone $\gg$ Dep $\gg$ *FinalHeavyσ $\gg$ Max

(15) /CVVL/ $\rightarrow$ CV[7]

	/púúr/	*2Tones-1μ	Max-Tone	Dep	*FinalHeavyσ	Max
a.	púúr				*!	
b.	púú				*!	*
c.	púr				*!	*
d.	púú.rú			*!		
e. $\rightarrow$	pú					**

[6]If the right generalization about distribution turns out to be syntactic, then this is phonologically derivable allomorphy sensitive to syntactic structure. This would mean that syntactic information must be available to the phonology, contra Strict Modularity but in line with other recent work (Simpson and Wu 2002; Dobler et al. 2011; Gribanova and Harizanov 2016; Richards 2016; Winchester 2016; Newell 2017; Branan 2018; Martinović 2019; Sande et al. 2026, a.o).

[7]Other possible candidates like, for example, [rú] would be ruled out by other high-ranking faithfulness constraints like Linearity, left out here for simplicity. Harmonically bounded candidates like [pá], which violates Dep and has three violations of Max, are left out.

(16) /CVL/ → CV

	/lõmbir/	*2Tones-1μ	Max-Tone	Dep	*FinalHeavyσ	Max
a.	lõmbir				*!	
b.	lõmbiri			*!		
c. →	lõmbi					*

- I assume that phonology has already applied to long forms before shortening (which must be a very late process; see discussion below) applies.
- This means, among other things, that /CVn/ words have the shape /CṼṼn/, where lengthening and nasality have productively applied, before shortening happens.
- This cyclicity simplifies the explanation of why vowel nasality persists on short-form vowels whose corresponding form has a coda nasal.

(17) /CṼṼn/ → CṼ

	/ɲɔ̃ɔ̃n/	*2Tones-1μ	Max-Tone	Dep	*FinalHeavyσ	Max
a.	ɲɔ̃ɔ̃n				*!	
b.	ɲɔ̃n				*!	*
c.	ɲɔ̃ɔ̃				*!	*
d.	ɲɔ̃ɔ̃nɔ̃			*!		
e. →	ɲɔ̃					**

- For words that already have a light final syllable, the *FinalHeavyσ constraint vacuously applies, and the faithful candidate, which violates none of the relevant constraints, is optimal.

(18) /CVɓ/ → CVɓ

	/k^{h}uɓ/	*2Tones-1μ	Max-Tone	Dep	*FinalHeavyσ	Max
a. →	k^{h}uɓ					
b.	k^{h}u					*!
c.	k^{h}uɓu			*!		

(19) /CV/ → CV

	/kókó/	*2Tones-1μ	Max-Tone	Dep	*FinalHeavyσ	Max
a. →	kókó					
b.	kó					*!*

- CVV and CVC words that bear two tones exceptionally do not show truncation to a light syllable.
  - This is due to the undominated constraints against two tones associated with a single mora, and requiring realization of input tones.
- In (20e) we finally see evidence that Dep must be ranked above *FinalHeavyσ, which until now has not been relevant.

(20) /CV̀V́/ → CV̀V́

	/pàá/	*2Tones-1μ	Max-Tone	Dep	*FinalHeavyσ	Max
a. →	pàá				*!	
b.	pà		*!			*
c.	pá		*!			*
d.	pă	*!				*
e.	pàtá			*!		

- The only remaining alternations in (12) not yet account for are the vowel quality alternations between singular and plural. These are of two types:
  i. Long-form diphthongs losing the first member of the diphthong in the corresponding short forms.
     - There is good reason to think that the second vowel in a diphthong is the target vowel, while the first is shorter, more transitional semivowel.
     - I assume that constraints prevent deletion of the target/longer portion of the diphthong, resulting in truncation to the second V rather than the first.
  ii. Some vowels show alternations before coda [n].
     - As previously mentioned, this is a language-wide phenomenon, seemingly due to phonotactic constraints on which vowel qualities can be both long and nasalized.
     - For example, we never see [ɛ̃ɛ̃n] sequences in the language; /ɛ/ raises to [ɪ] in a syllable with a nasal coda: [man.ɟɪ̋ɪ̋n] 'papaya', [man.ɟé.n-ɑ́] 'papayas'.
     - This nasal raising process must apply to roots before truncation applies, because the output of truncation is the raised, nasalized vowel without a following coda [n]: [man.ɟɪ́].

- **The timing of truncation:**
  - Truncation must happen late, for several reasons:
    - It operates on the output of other regular phonological operations:
      - Vowel raising before coda /n/, as just mentioned.
      - Vowel nasalization before a coda N, as in (17).
    - Its distribution depends on knowing where a given word sits within the structure of an entire noun phrase, as we will see in §5

- **Some implications of truncation in Lobi:**
  - This is truncation to a particular prosodic size, not subtraction of a particular amount of structure.
    - For many long/short alternations, we could say there's subtraction of a single segment, with regular phonological resolution of final diphthongs, etc., which would give the right outputs.
    - Subtraction cannot account for all of the forms, though:
      - It cannot straightforwardly account for why we don't see any segmental loss for final /CV/ or especially /CVɓ/ syllables (12r).
      - And a subtraction account cannot explain why we get loss of two segments, rather than just one, in CVVL words like /púúr/ 'pigeon' (12h).
    - A truncation account, on the other hand, accurately derives all surface alternations.
      - Lobi shows that truncation to a particular template is not always satisfied (contra the Satisfaction Condition of McCarthy and Prince (1996, 1999).
      - Modeling truncation via constraints is necessary in order to prevent truncation in the context of a higher-ranked markedness constraint.
  - Truncation typically involves shortening of an entire word (or stem)[8] to a particular prosodic size (Weeda 1992; Kurisu 2001; Inkelas 2014).
    - In Lobi, the size of the final *syllable* is restricted to satisfy a (violable) phonotactic constraint.
    - This data suggests that truncation may impose size restrictions on any prosodic unit, here the final syllable.

[8]See Downing (1997, 2000, 2006); Downing and Kadenge (2015, 2020) on sub-word stem size restrictions/requirements in Bantu. They argue that the Prosodic Stem is a level of the prosodic hierarchy, similar to the syllable or prosodic word, which we also now know can be targets of truncation.

5 Long and short distribution

The distribution of long and short forms in Lobi is predictable, and seems to be related to morphosyntactic structure rather than prosodic structure or the realization of a particular morpheme.

- Full transparency: I still do not know the best way to state (or formalize) the generalization about where long vs. short forms appear.
  - I will first show you the facts,
  - Then why I think the long/short alternation is not the realization of some morpheme.
  - I show you why I think this cannot be prosodic,
  - And then I'll point to two potential analyses determined directly or indirectly by morphosyntactic distribution. I'd love your feedback!

5.1 The distributional facts

Long forms of nouns appear…

- In isolation (21a,c),
- When followed only by a definite or demonstrative enclitic but no other modifier (21b),
- And when they are possessors (which precede the possessed noun) (21d).

(21) Long form distribution

a. cʊ̀ɔ̀r
house
'house'

b. cʊ̀ɔ̀r=rá/kɛ́/lé
house=DEF/DEM.PROX/DEM.DIST
'the/this/that house'

c. bɪ́ɪ́n
dog
'dog'

d. bɪ́ɪ́n=ǹ ˩nɔ́
dog-POSS leg
'a dog's leg'

Short forms of nouns appear…

- When the noun is modified by one or more following adjectives (22a,g),
- By a following numeral (22b),
- Or by a following quantifier or free-word demonstrative (22c);
- Or when preceded by a possessor (22d,e,h).

(22) Short form distribution

a. cɔ̀ sɪɛr
house red
'red house'

b. cɔ̀ bɪɛl
house one
'one house'

c. cɔ̀ ʊ́rɛ/úle
house DEM.PROX/DEM.DIST
'this/that house'

d. mɪ̂ɪ̂n ˩cɔ́
1SG.POSS house
'my house'

e. ɓɛɓɛ́=ɛ̀ ˩cɔ́
Kpekpe-POSS house
'Kpekpe's house'

f. cɔ sɛ j'ɛ́ɲɔ
house red two
'two red houses'

g. cɔ sɛ tɪ̂ɪ̂n
house.SHORT red.SHORT old
'an old red house'

h. mɪ̂ɪ̂n ˩cɔ́ sɛ
1SG.POSS house red
'my red house'

In noun compounds, the second noun is the head, and arguably the first noun is its complement. The first noun surfaces in its short form and the second in its long form.

(23) Short N1 and Long N2 in Noun Noun compounds

a. bɛ̂ nɔ́ɔ́
dog leg
'dog leg' (cf. [bɪ̂ɪ̂n] 'dog') (20230307_hls)

b. fa bíí
leaf soup
'leaf soup' (cf. [faa] 'leaf') (20230406_hls)

Adjective long forms appear…

- If there is no possessor and they are the final modifier in the noun phrase, cf. 'red' [sɪɛr] in (22a).

Adjective short forms appear…

- If there is a following full-word modifier in the noun phrase, cf. 'red' [sɛ] in (22f,g),
- Or if there is a possessor.[9]

To indicate intensity of an adjectival meaning, the adjective can be repeated two or more times (24) (Sande To Appeara).

(24) Adjective reduplication for intensity

a. cɔ pɔ́
house.SHORT clean
'a clean house'

b. cɔ pɔ́ pɔ́
house.SHORT clean clean
'a very clean house'

- If there is a modifier following the adjective, or a possessor modifying the noun phrase containing the adjective, all copies of the adjective surface as short (25).

[9]The facts are a bit more complicated than this, in that when there is a possessor and an adjective, which one surfaces as short varies with the interpretation, as elaborated on below.

(25) tʰr-a tɪ́ tɪ́ jɛ́ŋɔ
tree-PL old old two
'two old trees'

- If there is no higher modifier in the noun phrase, the final copy of the adjective surfaces as long, and the preceding copies as short (26).

(26) Adjectival intensity marked through repetition

a. tʰɪr tɪ́ɪ́n
tree old
'an old tree'

b. tʰɪr tɪ́ (tɪ́) tɪ́ɪ́n
tree old.SHORT old.SHORT old
'a very old tree'

5.2 The long/short distribution is not the realization of some morpheme

One analytical possibility is that a suffix in the long form or truncation in the short form is the realization of some morpheme.

- Initially, perhaps this seems plausible. If the short were basic and the long form was the realization of some morpheme high in the noun phrase, that would help to account for:
  - Why we only ever see one long form per noun phrase.
  - Why the one long form is always something syntactically high in the noun phrase, but not necessarily something that's always in the same linear position (e.g., initial vs final).
- This analysis does not work for several reasons, though:

Problem 1: First, we established in §4 that the long form is the basic, underlying form of a noun or adjective, and the short form is derived from it, not vice versa.

Problem 2: Second, while we see maximally one long form per noun phrase, we do not see one long form in every noun phrase. For example, in possessive constructions there are two noun phrases, but we only see one long form, as in (27a).

- If truncation is the exponent of some morpheme instead, then we get rid of Problem 1, above. However, Problem 2 is only exacerbated.
  - We can see many many (potentially infinite?) short forms per noun phrase: all but maximally one word in the noun phrase is short.
  - If each instance of truncation is the exponent of some morpheme, then there are potentially infinite instances of that morpheme per noun phrase, which just happens to correspond exactly with the number of nouns and adjectives that are not the single long form.
  - What would that morpheme mean?
- It does not seem to be the case that short or long forms are the realization of some morpheme.

5.3 The long/short distribution is not prosodically determined

The long vs short distribution cannot be defined prosodically:

- One possible explanation for conditioning truncation (the short form) would be to point to a prosodic boundary, or lack of prosodic boundary, in all cases where we see short forms as opposed to long forms.
  - For example, we might expect to find truncation only finally within a phrase, or only non-finally.
  - However, this does not match the data.
- We cannot say that long forms are only possible in phonological phrase-final position, because while most contexts where long forms appear are phrase-final (cf. 'house/dog' in (11a,b,c) and 'red' in (22a)) and most short-form contexts are non-final (22a,b,c,f), possessive contexts are an exception to this generalization:
  - Possessors precede possessees and surface in their long form, while possessees are short.
    - Even if possessors and possessees are argued to be in distinct phonological phrases, we would still expect the final word of the possessee, which is final in the whole noun phrase, to be long if heavy syllable are permitted phrase-finally, but it is not (27a,c).
    - Possessives and compounds are both contexts with adjacent nouns. If prosody were at play, we might expect them to show the same long/short properties, but they don't (27a,b)

(27) Prosody doesn't determine long vs short forms:

a. bíín=ǹ ꜜnɔ́
dog-POSS leg
'a dog's leg' (20230307_hls)

b. bɛ́ ꜜnɔ́ɔ́
dog leg
'dog leg' (20230307_hls)

c. míìn ꜜbɛ́
1sg.poss dog
'my dog' (20230407_ymy)

  - Additionally, there is variation in whether the final modifier in a possessed noun phrase surfaces as long or short (28).
    - This seems to come with a difference in meaning.
    - The short form of the final modifier of the possessed noun in the context of a possessor seems to be the default, discourse-neutral option, and the other seems exceptional, but I do not fully understand their distributions yet. For future work!

(28) a. míìn ꜜdú sɪɛr
1SG.POSS room red
'my red room' (Situation: I have a lot of rooms, and I'm talking specifically about the red room of mine)

b. míìn ꜜdú sɛ
1SG.POSS room red
'my red room'

c. míìn cɔ́ tɪ́ tíín
1SG.POSS house old old
'my old house'

d. mîîn cɔ́ tɪ́ tɪ́
1SG.POSS house old old
'my old house' (Hien offered that this would be natural in the context 'This is my old house.')

- I conclude that reference to prosodic hierarchy does not help to predict the correct distribution of long and short forms.

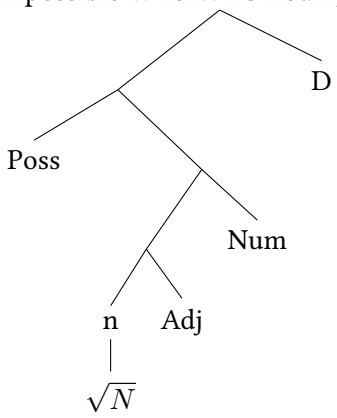
So what's going on? The next two sections present two possible ideas, both of which seem better to me than the ones presented so far, but both of which come with their own challenges for formalization.

5.4 Is the long/short distribution morphosyntactic?

I have been thinking of the single long form per noun phrase as being the syntactically (hierarchically) highest word in the noun phrase.

- One complication: This only applies to full words. Clitics are outside the calculation of what is "highest".
- Another complication: How are we defining what is syntactically "highest"?
  - Maybe something like this: The noun or modifier that is last to be merged into the syntactic tree is long. Other items in the noun phrase are short.

(29) A possible WYSIWYG noun phrase structure:



- Arguably, due to various scope and distribution reasons, D (where definite and demonstrative heads presumably live) is higher than Poss.
- Among D heads, full-word demonstratives affect long/short distribution, but definite and demonstrative clitics do not.
  - This suggests that we must differentiate between prosodic words and clitics, as well as syntactic height, in our generalization and formalization: *Only the hierarchically highest prosodic word is long.*
  - This requires our phonological constraints to refer to phonological information, prosodic structure, and syntactic structure.
- A potential problem: Is the relative height of Adj and Poss distinct in the examples repeated below? Currently unclear.

(30) a. mîîn ↓dú sɪɛr
1SG.POSS room red
'my red room'

b. mí̂ín ˡdú sɛ
1SG.POSS room red
'my red room'

Open questions:

i. If the generalization is truly syntactic, this complicates our constraint-based analysis of truncation:
   - Our constraints have to reference syntactic height, contra Strict Modularity.
   - If we use syntactic phases to calculate the domain of long/short distribution, what is the relevant phase? Maybe he full DP phase including the D head, but…
     * This is complicated, because there are presumably phases internal to this domain (like complex possessor DPs), but the internal structure of possessors factors into the overall calculation of height.

   (31) Internally complex possessors
   a. [kũ gbáán=ǹ] ˡlɔ́
   man tall=POSS farm
   'A tall man's farm' (cf. [kuun] 'man'; [gbáán] 'tall; [lɔ́ɔ́] 'farm')
   b. [kʰɛ́r=ˡrá=à] ˡlɔ́
   woman=DEF=POSS farm
   'The woman's farm'

ii. Do we see other kinds of phonological alternations sensitive to syntactic height?
   - Maybe. See, for example, phonological accounts that adopt constraints referring to C-Command, phasehood, and other syntactic structural notions (McPherson and Heath 2016; Rolle 2018; Fenger and Weisser 2025, a.o).
   - Also nuclear stress; see below.

iii. What other kinds of data would be useful for testing the structure of DPs in Lobi?

5.5 A radical, half-baked idea: Long form retention as nuclear stress

One final analytical possibility I have recently been considering is whether something similar to nuclear stress within the noun phrase is at play.

- How this would work:
  - The long form is retained only in the "stressed" word.
  - Truncation targets all words without nuclear stress.
- Reasons for thinking this:
  - It explains why there is maximally one long form (=nuclear stress) in every noun phrase.
  - The distribution of the long form is similar to the distribution of English nuclear stress:
    * The long form (=nuclear stress) can be a focused item (as in the difference between 'my red room' and 'my RED room', where red is exceptionally long in the latter) or it can be the phrase-final stressable thing.[10]

[10]There are several competing analyses and unclear judgments about nuclear stress (cf. Halle and Vergnaud 1987; Cinque 1993; Kahnemuyipour 2009 in English, most of which do not actually dig into the details of noun phrase stress, but these principles are generally agreed upon.

(32) English nuclear stress in noun phrases

a. A very old trée.

b. My red róom.

c. Mý red room.

- A potential caveat: Why would the prosodic restriction only apply to nouns and adjectives, not possessive pronouns, numerals, or demonstratives?

(33) Numerals do not shorten

a. bɛ̃-nã nʊ̀ɔr
dog-PL ten
'ten dogs'

b. bɛ̃-nã nʊ̀ɔr/*nɔ j'ɛ́rɛ
dog-PL ten DEM.PROX.PL
'these ten dogs'

- Because truncation is phonological as argued in §4, any account will have to deal with this problem.

6 Implications and conclusions

- This talk has described the form and distribution of Lobi stem alternations among nouns and adjectives.
- **Alternations in form**:
  - I have shown that the long/short alternations in form are neither suppletive (contra Sande (To Appeara)) nor additive (contra the typical assumption for Gur languages Miehe and Winkelmann (2007), which has also explicitly been stated for Lobi (Becuwe 1982; Kambou 2003)).
  - Instead, short forms are predictably derived from long forms via truncation of the final syllable.
  - Truncation must be derived via phonological constraint interaction, because it can be blocked exactly in the case that truncation would violate a higher ranked markedness constraint.
- **Distribution:** Where we stand on the distribution of long/short forms:
  - The long/short alternation is not the realization of some morpheme.
  - The long/short distribution is not determined by position within a given prosodic domain.
  - The long/short distribution seems related to syntactic structure, either directly or indirectly via something like a nuclear stress rule (which itself makes reference to syntactic structure).

  Open questions:
  - What kind of data would differentiate the two plausible analyses?
  - How to implement the fact that truncation only applies to nouns and adjectives, not other categories?
- A final thought: Maybe the right way to think of it is that Lobi truncation, and nuclear stress, are both phonological processes that directly references syntactic structure.
- If this is on the right track, then our phonological constraints need to be able to reference the fine details of syntactic structure, with implications for modularity.

7 References

Anderson, Stephen R. 1992. *A-morphous morphology*. 62. Cambridge University Press.

Becuwe, Jacques. 1982. Eléments de phonologie et de grammaire du lobiri (parler de Bouna, Cote-d'Ivoire). Doctoral Dissertation, Université de la Sorbonne Nouvelle (Paris 3).

Branan, Kenyon. 2018. Relationship preservation. Doctoral Dissertation, Massachusetts Institute of Technology.

Caballero, Gabriela. 2008. *Choguita Rarámuri (Tarahumara) phonology and morphology*. University of California, Berkeley.

Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic inquiry* 24:239–297.

Clem, Emily Catherine. 2019. Agreement, case, and switch-reference in Amahuaca. Doctoral Dissertation, University of California, Berkeley.

Dobler, Eva, Heather Newell, Máire Noonan, Glyne Piggott, Mina Sugimura, Lisa deMena Travis, and Tobin Skinner. 2011. Narrow syntactic movement after spell-out. Ms., McGill University.

Downing, Laura J. 1997. Prosodic misalignment and reduplication. In *Yearbook of morphology 1997*, 83–120. Springer.

Downing, Laura J. 2000. Morphological and prosodic constraints on Kinande verbal reduplication. *Phonology* 17:1–38.

Downing, Laura J. 2006. *Canonical forms in prosodic morphology*. 12. Oxford University Press.

Downing, Laura J, and Maxwell Kadenge. 2015. Prosodic stems in Zezuru Shona. *Southern African Linguistics and Applied Language Studies* 33:291–305.

Downing, Laura J, and Maxwell Kadenge. 2020. Re-placing PStem in the prosodic hierarchy. *The Linguistic Review* 37:433–461.

Eberhard, David M., Gary F. Simons, , and Alison J. Robinson. 2026. Ethnologue: Languages of the world. URL `Dallas, Texas: SIL Global. Online version: https://www.ethnologue.com/language/lob/`.

Fenger, Paula, and Philipp Weisser. 2025. Matching domains: The syntax, morphology, and phonology of the verb in Sinhala. *Natural Language & Linguistic Theory* 43:2523–2582.

Gribanova, Vera, and Boris Harizanov. 2016. Locality and directionality in inward-sensitive allomorphy: Russian and Bulgarian. *The Morphosyntax-Phonology Connection* .

Halle, Morris, and Jean-Roger Vergnaud. 1987. Stress and the cycle. *Linguistic inquiry* 18:45–84.

Inkelas, Sharon. 2014. *The interplay of morphology and phonology*, volume 8. Oxford University Press.

Kahnemuyipour, Arsalan. 2009. *The syntax of sentential stress*. Oxford: Oxford University Press.

Kambou, Moses Kwadwo. 2003. Morphophonology of verbo-nominal roots in lobire. *Journal of Dagaare Studies* 3:1–20.

Kurisu, Kazutaka. 2001. *The phonology of morpheme realization*. University of California, Santa Cruz.

Labouret, Henri. 1958. Notes grammaticales sur le lobi. In *Nouvelles notes sur les tribus du rameau lobi: leurs migrations, leur évolution, leurs parlers et ceux de leurs voisins*, volume 54 of *Mémoires de l'IFAN*, 101–125, 178–209, 210–275. Institut Français de l'Afrique Noire.

Lamothe, Charles PB. 1964. *Esquisse du système verbal lobi*. CNRS ua.

Lamothe, Charles PB. 1966. *Esquisse du système grammatical lobi*. CNRS ua.

Martinović, Martina. 2019. Interleaving syntax and postsyntax: Spellout before syntactic movement. *Syntax* 22:378–418.

McCarthy, John, and Alan Prince. 1996. *Prosodic morphology 1986*. Rutgers University Center for Cognitive Science.

McCarthy, John J, and Alan Prince. 1999. Prosodic morphology (1986). In *Phonological theory: The essential readings*, ed. John Goldsmith, 238–288. Malden, MA: Blackwell.

McPherson, Laura, and Jeffrey Heath. 2016. Phrasal grammatical tone in the Dogon languages. *Natural Language & Linguistic Theory* 34:593–639.

Miehe, Gudrun, Ulrich Kleinewillinghöfer, Manfred von Roncador, and Kerstin Winkelmann. 2012. Overview of noun classes in Gur (ii)(revised and enlarged version). *Noun class systems in Gur languages* 2:5–37.

Miehe, Gudrun, and Florian Tham. 2007. Lobiri. In *Noun class systems in Gur languages*, volume 1. Rüdiger Köpper Verlag.

Miehe, Gudrun, and Kerstin Winkelmann. 2007. *Noun class systems in Gur languages*, volume 1, Southwestern Gur

Languages (without Gurunsi) of *Gur Monographs*. Rüdiger Köpper Verlag.

Newell, Heather. 2017. Nested phase interpretation and the PIC. In *The structure of words at the interfaces*, ed. Heather Newell, Máire Noonan, Glyne Piggott, and Lisa deMena Travis, 20–40. Oxford: Oxford University Press.

Richards, Norvin. 2016. *Contiguity theory*. MIT press.

Rolle, Nicholas. 2018. Grammatical tone: Typology, theory, and functional load. Doctoral Dissertation, UC Berkeley.

Sande, Hannah. To Appeara. Adjectival reduplication in Lobi is morphosyntactic. In *ACAL 55 proceedings*. Language Science Press.

Sande, Hannah. To Appearb. The morpheme-specific behavior of floating tones in Lobi (Gur). In *ACAL 56 proceedings*. Language Science Press.

Sande, Hannah, Emily Clem, and Maksymilian Dabkowski. 2026. Discontinuous vowel harmony in Guébie: Cyclic interleaving of syntax and phonology. *Language* 102:2:249–280.

Sande, Hannah, and Marcus Ewert. 2026. Twisted Tongues: An online database management tool for linguistic fieldworkers. URL `https://twisted-tongues.com`, accessed: 2026-11-10.

Sande, Hannah, and Sansan Claude Hien. To Appear. The puzzling distribution of downstep in Lobi (Gur/Mabia). In *New perspectives on phonological domains*, ed. Eirini Apostolopoulou, Maksymilian Dabkowski, Martin Krämer, and Hannah Sande, Topics in phonological diversity. Language Science Press.

Sande, Hannah, and Dylan Worden. 2026. A synchronic scalar vowel shift in Lobi (Gur/Mabia) plural formation. Paper presented at the Annual Meeting on Phonology (AMP 2026), University of Maryland, October 9–11, 2026.

Simpson, Andrew, and Zoe Wu. 2002. IP-raising, tone sandhi and the creation of S-final particles: Evidence for cyclic spell-out. *Journal of East Asian Linguistics* 11:67–99.

Trommer, Jochen, and Eva Zimmermann. 2014. Generalised mora affixation and quantity-manipulating morphology. *Phonology* 31:463–510.

Urbanczyk, Suzanne. 2006. Reduplicative form and the root-affix asymmetry. *Natural Language & Linguistic Theory* 24:179–240.

Urbanczyk, Suzanne Claire. 1996. Patterns of reduplication in Lushootseed. Doctoral Dissertation, University of Massachusetts, Amherst.

Vaillant, Par M. 1967. Esquisse grammaticale du lobiri. Doctoral Dissertation, Universite de Dakar.

Weeda, Donald Stanton. 1992. Word truncation in prosodic morphology. Doctoral Dissertation, The University of Texas at Austin.

Winchester, Lindley. 2016. Morphosyntactic features and contextual allomorphy: Evidence from Modern Standard Arabic. *NELS 47* .

Zimmermann, Eva. 2017. *Morphological length and prosodically defective morphemes*. Oxford University Press.