Stress and Intonation in Sereer-Sine: A Preliminary Analysis

Raimundo Cox-Casals
McGill University
Linguistics
raimundo.coxcasals@mail.mcgill.ca

William Melançon
McGill University
Linguistics
william.melancon@mail.mcgill.ca

Abstract

In this preliminary study, we present an overview of the stress patterns and intonational system of Sereer. Through phonetic analysis and statistical tests of data recorded from a speaker, we identify Sereer as a stress language, where stress is primarily cued by syllable intensity rather than pitch. We present a metrical analysis of stress and a structural analysis of intonation. Crucially, stress and intonation are demonstrated to operate independently, with stress relying on amplitude and intonation governed by f0 patterns. This study contributes to the understanding of Sereer's prosodic systems and may aid in future morphological approaches to the language.

1 Introduction

Sereer (also called Sereer-Sine or Sine-Saloum) is a Senegambian language of the Atlantic branch of Niger-Congo (Becker, 1985). It is primarily spoken in Senegal's modern-day regions of Fatick, Kaolack, and Kaffrine, corresponding to the historical Sine and Saloum Kingdoms.[1] Migration has expanded its use to regions such as Thiès and Diourbel, where another language, Sereer-Noon, is also spoken (Merrill, 2023). While tonal Niger-Congo languages in the region have received significant phonological attention, Sereer's prosodic features remain understudied. Sereer is non-tonal by McCawley (1978)'s definition, but it exhibits a stress system primarily marked by syllable intensity. Interestingly, the pitch cue is taken over by intonational features and is almost completely unconcerned by the location of word-level stress, unlike in other languages like English (Lieberman, 1960) where f0, amplitude, and duration combine to mark stress. While in Sereer, peaks in intensity are used to signal stress, and peaks and drops in pitch are used to signal (H)igh and (L)ow intonational targets, these two phonetic aspects mostly behave independently from each other. Figure 1 is a pitch tracker and intensity tracker of for the word /foːfiʔ/ 'water' when it is aligned with the right edge of an utterance-medial phrase. We visualize that the pitch on the second syllable [fiʔ] hits an H target to mark the end of a prosodic domain, a phonological phrase (PPh) in medial position in this case (surrounded by two other PPhs on either side). However, the intensity on the first syllable [foː] is higher than the second syllable to mark the stress. The two dimensions appear to work inversly from each other in this light.

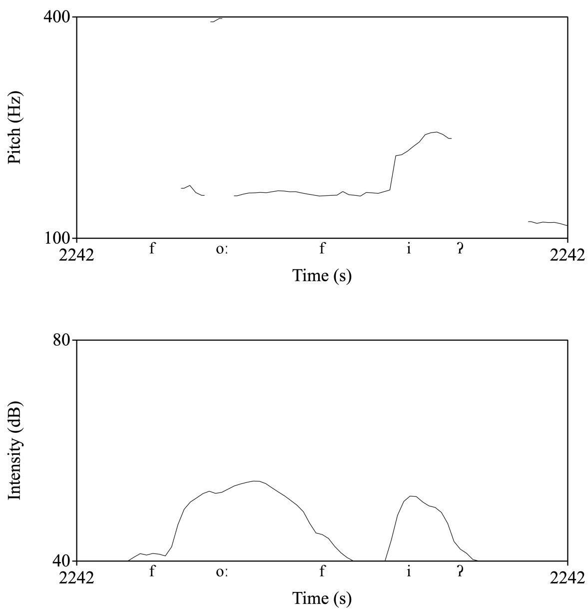

Figure 1: Pitch and intensity graphs for /foːfiʔ/

Our first task is to identify stress, find its location within words, and show how it is cued by amplitude. This was done through a phonetic and statistical analysis of recorded, elicited data from

[1]Sereer has 1,810,000 speakers in Senegal (2021 census)

a native speaker from the city of Thiès. In §3 we outline how we tested for the presence of stress and §4 presents evidence for our conclusion that Sereer is a stress language.

Once stress is demonstrated, we proceed to propose the phonological parameters of stress assignment from a selected list of words and their surface stress patterns. Sereer is an SVO language with a rich system of verbal morphology which can attach functional information before and after the verb stem. Because of the current uncertainty about morphosyntactic categories surrounding verbal morphology (i.e. whether verbs take affixes, clitics, conjugation, etc.), we cannot definitively label any functional morphemes at this stage, and thus cannot diagnose the contribution of different morphological categories to the stress domain.[2] To get around this issue, we will look at proper names and monomorphemic nouns as their minimal morphosyntactic structure transparently displays the default mechanics and parameters of stress assignment. Like in English, nouns and verbs may also have differing stress rules so identifying which elements fall inside the stress domain of the verb or which elements come with prespecified stress may prove difficult. What we find is that verb stems prefer to be prominent and are assigned stress, while functional morphemes adjoined to the verb stem do not all carry stress (though some do). As for the default system which applies to nouns, in the first part of §5 we show that Sereer builds non-iterative trochaic feet from right-to-left, and that codas contribute moraic weight (though syllable-final consonants may not).

Next, in the second half of §5, we will give a detailed preliminary analysis of the intonational system of Sereer and how it does not interact with stress. Sentences of the same recorded data will be presented with an f0 tracker for identifying pitch contours to accompany the intonation analyses. For simplicity, we only look at declarative sentences with a maximum structure of subject, verb, direct object, indirect object (in this order). Selected sentences will have a maximum subdivision of three PPhs along the utterance, and a minimum of one PPh encompassing the entire utterance. We find that prosodic boundaries are signaled by phrase and boundary tones, pauses, and glottal stop insertion when a phrase is vowel-initial. Utterances can be subdivided differently depending on speech rate but always follow the same rule of splitting between syntactically relevant constituents, e.g. dividing the subject and predicate into separate PPhs.

For a final note on intonation, utterance final PPhs in Sereer tend to promptly drop towards a final L% after the syllable with an H tone at the left edge of the phrase. If the PPh is more than two syllables, there is a drop after the H at the left edge and a second audible drop on the penultimate syllable which anticipates the final L% on the last syllable. It appears that the final L% infects the syllable immediately to its left and we will show that it is not concerned with the location of stress or morphosyntactic edges, further suggesting the independence of the two systems.

To conclude in §6, we summarize: 1) the mechanics of stress and its cueing by intensity, 2) the intonation system in declaratives and its cueing by pitch, and 3) the independence of these two prosodic features. We also propose directions for future research in Sereer's morphosyntax and phonology.

2 Theoretical Background

2.1 Defining Prosodic Structure

Many theories of prosodic organization have been proposed in the literature. In Selkirk (1986)'s view, prosodic structure is a hierarchy of purely prosodic constituents. This view, adopted here and by scholars such as Ferraira (1993), assumes that branches in the prosodic tree cannot cross, ensuring that lower-level constituents belong to only one higher-level constituent. The highest node in the structure is the Utterance (U) as it constitutes the entire tree. Directly below is the Intonational Phrase (IP) which tends to end with a pitch drop for a declarative and normally followed by a pitch reset (Nespor & Vogel, 2007). Below the IP is the Phonological Phrase (PPh or Φ) which, like Selkirk, we will not distinguish with the intermediate phrase (ip) or clitic group (C). The PPh can end with a phrase tone and sometimes a caesura. Selkirk (1986) and Ferraira (1993) talk about an algorithm which scans the syntactic structure of a sentence for the right boundary of a phrasal constituent to align it with the PPh boundary. This is the level most connected to syntax. Below the PPh is the Phonological or Prosodic Word (PWd or

[2] We will not discuss how morphosyntactic structure relates to verbal morphology that attaches phonologically close to the verb to enter the stress domain or not.

$$\begin{array}{c} \mathrm{U} \\ | \\ \mathrm{IP} \\ | \\ \Phi \\ | \\ \omega \\ | \\ \mathrm{Ft} \\ | \\ \sigma \\ | \\ \mu \end{array}$$

Figure 2: Prosodic hierarchy

ω) which corresponds to the distinction between content and function words. Selkirk solely considers content words to be PWds on their own since they are the domain for stress assignment, and that function words like clitics attach to lexical hosts. Thus, the PWd corresponds to a content morpheme with attached functional morphemes. The PWd is in turn composed of Feet (Ft or φ), Syllables (σ) and Moras (μ) which we will revisit in the following section on stress.

2.2 Defining Stress

As a phonological system, stress is a marking of relative prominence placed on syllables within words. There is a structural aspect to stress as it involves the hierarchical organization of constituents, as well as a metrical aspect as it aligns with a metrical grid (Liberman & Prince, 1977). At the phonological word level, Hayes (1980) observed that stress placement is often sensitive to syllable weight. This concept was expanded in Moraic Theory (Hayes, 1989), which posits that light syllables, of the shape C_0V, have a syllable-internal structure containing one mora (μ), and heavy syllables, of the shape C_0VV or C_0VC, contain two moras. Stress is typically drawn to heavy syllables. It should be noted that coda consonants do not contribute a mora in all languages. In some languages, CVC syllables have no weight value applied to the final C and are thus monomoraic, patterning with CV syllables (Goldsmith, 1990).

In this paper we will notate monomoraic (light) syllables as σ_μ and bimoraic (heavy) syllables as $\sigma_{\mu\mu}$ encoding their mora count. We adopt a parametric theory of stress assignment taken from Prince (1983) and Hayes (1980) wherein stress is predictable from a given set of binary parameters that scan phonological words and parse syllables into feet. The parameters are listed below:

1. Foot type: whether a foot is right-headed (iambic) or left-headed (trochaic)
2. Foot size: whether a foot has two integrants (bounded) or can have more than two (unbounded)
3. Foot binarity: whether a foot can be subminimal (degenerate) or must be binary
4. Weight-by-Position: whether coda consonants are assigned a mora
5. Directionality: starting from which edge of the word are feet parsed
6. Iterativity: whether identical feet are subsequently parsed after the initial foot
7. End rule: whether the leftmost or rightmost foot recieves primary stress

Altogether, these binary specifications combine to syntagmatically locate syllables inside words and assign them to a strong metrical position. In §5 we will review different parameter settings and discuss how well they predict the surface realizations of stress.

As a phonetic event, prominence can be signaled through an f0 fluctuation, a rise in amplitude, and lengthening of the stressed syllable. For Sereer, we will show in §4 that this corresponds to amplitude. Primary stress will be marked with an acute accent (á) and secondary stress with a grave accent (à) on the vowel in the nucleus of the stressed syllable. In grid theoretical diagrams, we will mark the location of stress with ‘x’ above the syllable, feet with parentheses ( ), and extrametrical syllables with angle brackets ⟨⟩.

2.3 Defining Intonation

The pitch modulation over Sereer utterances is not specific to a certain syllable of a word, but rather is a property of the sentence. “Intonation [...] refers to the use of *suprasegmental* phonetic features to convey ‘postlexical’ or *sentence-level* pragmatic meanings in a *linguistically structured way*” (Ladd, 2008). Suprasegmental features include aspects of the fundamental frequency (f0), intensity, and duration. We will use the terms pitch and f0 interchangeably though one refers to the psychological interpretation and the other a physical property. Some disagreement stands for which term is a more appropriate identification of suprasegmental material but here we will avoid the issue of distinguishing them. We will also use the terms intensity, loudness and amplitude interchangeably. The lexically contrastive suprasegmentals are either stored in the lexicon or assigned at an earlier stage of the phonological derivation. Below are some examples.

1. Sereer vowel length distinctions
   short [muɲ] 'endurance'
   long [muːɲ] 'toothy grin'
2. Catalan stress distinctions
   strong-weak [ˈkala] 'cove'
   weak-strong [kaˈla] 'to peirce'
3. Japanese pitch accent distinctions
   LH [aꜛme] 'rain'
   HL [aꜜme] 'candy'
4. Thai tone distinctions
   mid tone [kʰāo] 'fishy'
   falling tone [kʰâo] 'rice'

These minimal pairs consist of the same segmental material but their meanings are distinguished by the location of stress, pitch changes, or the identity of tones. Such suprasegmental material which is important for signaling lexical contrasts is reliant on the same phonetic cues as intonation. The difference is that intonation operates at a post-lexical sentence level where it conveys meaning that applies to whole utterances. Intonation can indicate questions, focused information, discourse functions and flow, speaker's identity and attitude, etc. Depending on the language, these aspects can be explicitly marked or indicated by prosody. Two things here are crucial: 1) intonational or post-lexical prosodic entities like boundary tones, caesuras, and accents are on a more independent tier than the lexical prosodic features, and 2) all of these features interact along the utterance. As we will see, this interaction is less evident in the case of Sereer since intonational tones do not seem to be anchored by stressed syllables. For analyzing intonation, we use Ladd (2008)'s Autosegmental-Metrical approach, where intonational features are expressed as a series of non-grammatical tones. We use the notational convention of Tones and Break Indices of Pierrehumbert (1980) where, unlike the tones and pitch accents shown in the Thai example above, H and L represent tonal targets that are not lexical, but intonational. Herein, we will refer to these intonational features as tones (T), not to be confused with lexical tones which Sereer lacks. Below is a list of the types of tones with their functions:

1. Boundary tone (T%): realized at the edges of intonational phrases to mark the start or end of the domain. For declaratives, the left edge of the utterance is marked with a H% and the right edge with a L%.

2. Pitch accent (T*): realized on words that carry the most information in a domain (H*) or previously mentioned information (L*).

3. Phrase accent (T or T-): realized as transitions on syllables between pitch accents and boundary tones, and used to mark the edges of intermediate phrases.

These three tones will be used to label Sereer utterances when discussing their intonational pitch contours. Epenthetic glottal stops in phonetic transcriptions will be written in brackets at the start of the vowel-initial word that they affect. Caesuras (pauses) in phonetic transcriptions will be written with two pipes (‖) when long, and one pipe (|) when short.

3 Methods

Data was collected using a ZOOM H4n external microphone over two sessions conducted in December, 2024 with the native speaker of the Sereer-Sine variety, from the city Thiès, Senegal. Various simple SVO sentences were elicited, as we do not yet have a deep understanding of Sereer morphosyntax, and need to be confident about prosodic structure and edge-alignment. Most sentences were elicited more than once, at a slow speed where the speaker tended to separate the subject, verb, and object into 3 PPhs, and at a regular conversational speed, where the speaker tended to group the verb and object together into one PPh (sometimes subject and verb were grouped). The utterances containing the subject, verb, and object respectively were then aggregated into two Excel sheets. The CV profiles, pitch, and intensity values for each of their syllables were analyzed and collected on Praat. The prosodic environments of each word (e.g. left- or right-alignment within medial phrases, or utterance initial/final position) were also recorded. The sheets were then exported into a CSV file, on which several statistical tests were performed using a Python script.

3.1 Relationship Between Pitch and Intensity

For every syllable in the dataset, pitch and intensity were recorded as an (x, y) pair respectively on a scatter plot. The Pearson Correlation Coefficient was then calculated, to determine how strong the correlation between the two variables is. A linear regression was fitted to the data for visualization. The Standard Error of Estimate was then calculated to evaluate how well the linear regression fit the data. Finally, both the regression line and points were plotted. This test was done to determine whether pitch and intensity interact in the signaling of stress or intonation. If pitch is not a significant cue for stress, then it is solely being controlled for intonation and intensity is responsible for cueing stress, as illustrated by Figure 1.

3.2 Locating Stress

A list of each phonological word that appears at least once in the dataset was collected. Then, for each word, the mean intensity of each of its respective syllables was calculated, and its strongest syllable was labeled as stressed.

3.3 Intensity as a Correlate of Stress

All stressed syllables were aggregated, and mean intensity and standard deviation were calculated. This process was repeated for unstressed syllables, and then an independent samples t-test was conducted between both sets to determine whether their difference in intensity is statistically significant. Cohen's *d* and confidence intervals were subsequently calculated to evaluate how meaningful the difference is between both groups and the certainty of the result.

3.4 Pitch Contours in Final Phrases

Given our observation that pitch in final phrases tends to drop towards an L% tone earlier than in English (i.e. not on the final syllable), we tested the pitch contours of these. Every utterance-final PPh was collected and sorted into 1 of 3 subsets: 2-syllable, 3-syllable, or 4-syllable phrases. Final PPhs containing 5 syllables or more were not found in the data. For each of these subsets, mean frequency and standard error were calculated for each syllable, and plotted on a syllable vs. mean frequency plot with error bars. These graphs will be presented and discussed in §5.2.5.

4 Results[3]

4.1 Relationship Between Pitch and Intensity

The Pearson Correlation Coefficient indicated a positive, relationship of little strength at $r(1131) = .091$, $p = .0029$, with an $SEE = 175.2$. The total sample size of our dataset was n = 1139 syllables, but 8 had to be excluded due to Praat not being able to read their pitch (they were kept in the dataset for intensity calculations). The very low PCC and quite high SEE for our range seem to confirm our hypothesis that pitch and intensity are unrelated. Since p-value $< .05$, we know that these values are statistically significant; in other words, there is some positive correlation between pitch and intensity, but it is very weak, and it is highly likely that the main cues for these two are different but still interact with one another in certain cases.

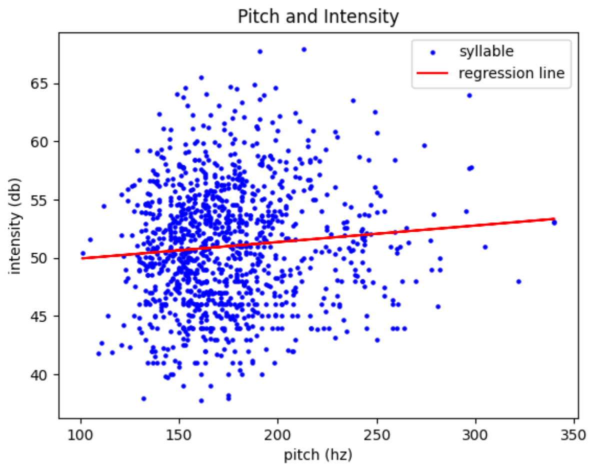

Figure 3: Correlation between pitch and intensity

4.2 Intensity as a Correlate of Stress

The independent samples t-test had stressed syllables ($n = 65$) with a mean of $M = 52.81$ ($SD = 4.20$) and unstressed syllables ($n = 126$) with a mean of $M = 50.65$ ($SD = 3.90$). The results of the test gave $t(189) = 3.537$, $p < .05$, Cohen's $d = 0.540$, 95% CI [0.956, 3.367]. Generally, a t-value larger than 2 means there is a significant difference between the two groups (larger than variation within

[3] Average intensities were not included in this section, as we did not directly use or interpret that data in any meaningful way. Rather, we used it to calculate the strength of the correlation between intensity and stress; see §4.2

a group), and the low p-value confirms that this is statistically significant. Further, Cohen's d hovering around $\tilde{0}.5$ indicates a moderate effect size, meaning stressed and unstressed syllables do not differ wildly in their intensity but the difference is certainly identifiable. Thus, we conclude that intensity is indeed the main cue for stress, and that the difference in intensity between stressed and unstressed syllables is statistically sound.

4.3 Pitch Contours in Final Phrases

Graphs for average pitch per syllable of 2-syllable, 3-syllable, and 4-syllable words respectively are given by Figures 4, 5, 6. The exact values for mean frequency and standard error are given in the following tables:

Syllables	σ_1	σ_2	σ_3	σ_4
2	177.32	139.46	$\emptyset$	$\emptyset$
3	194.78	160.26	137.26	$\emptyset$
4	189.43	172.00	164.21	151.00

Table 1: Mean Frequency (Hz) per Syllable

Syllables	σ_1	σ_2	σ_3	σ_4
2	3.57	2.35	$\emptyset$	$\emptyset$
3	3.14	4.57	2.18	$\emptyset$
4	6.26	7.56	4.68	2.89

Table 2: Standard Error per Syllable

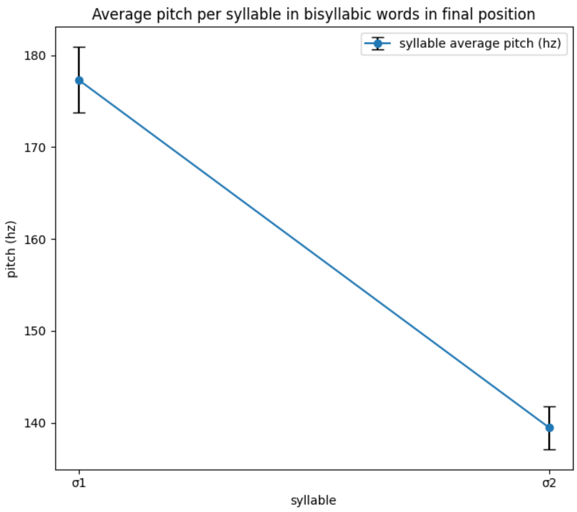

Figure 4: Average pitch differences in two-syllable utterance-final phrases.

As expected, words in final position aligned with edges exhibit an H tone at the left edge, and

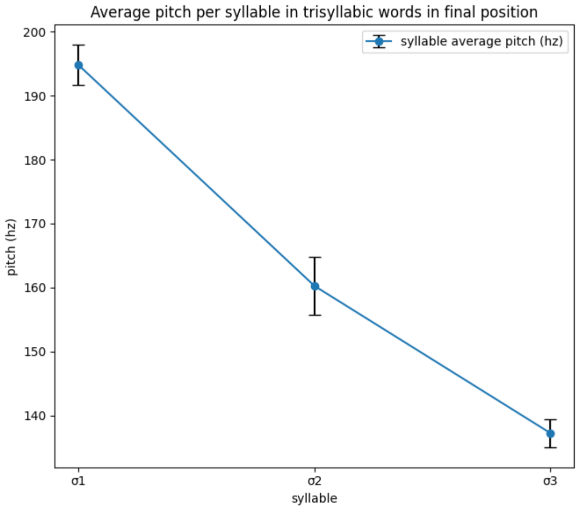

Figure 5: Average pitch differences in three-syllable utterance-final phrases.

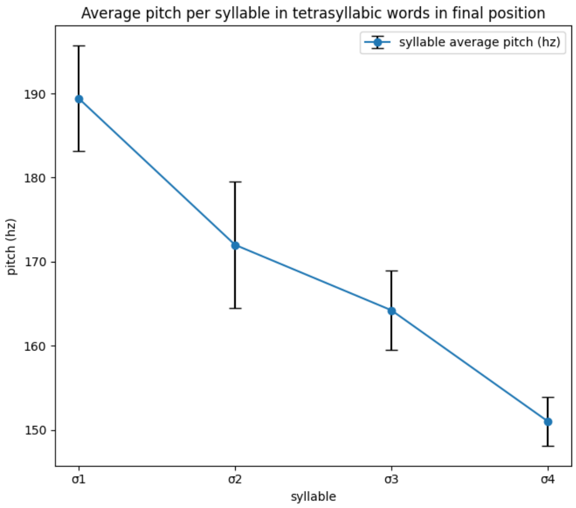

Figure 6: Average pitch differences in four-syllable utterance-final phrases.

an L% boundary tone on the final syllable. Further, in PPhs with more than 2 syllables, pitch immediately decreases after the first syllable (cf. Figures 11 and 14). Lastly, the low value of the standard errors show that this pattern is virtually universal, as $\sigma_1 - \max SE_{\sigma_1} > \sigma_2 + \max SE_{\sigma_2}$.

5 Discussion

5.1 Parameters of Stress

Now that we have a grasp on which syllables in our data are stressed, we will build a theory for the parameters that Sereer observes when allocating stress. It should be noted that in our results, word-initial syllables were always strengthened as they are in many languages, though this does not necessarily entail metrical stress. As we will see, when longer words have primary stress somewhere in the middle, it is difficult to tell if the prominence heard on the initial syllable is an effect of rhythmicity. In other words, if iterative feet create a rippling effect of secondary stresses along the word after the primary stress, or if there is only one stress per word and initial prominence is a consequence of left-edgedness. We will primarily be looking at names and hypocoristics, but will also briefly touch on verbs.

We know from short two-syllable names, such as [báda] 'Bada' and [sáli] 'Sali', that Sereer parses trochees. When one of the two syllables is heavy, like with [fáːtu] 'Fatou' and [ɟégaːn] 'Diégane', stress is attracted to the heavy syllable.

(1)

x	x
$(\sigma_\mu\ \sigma_\mu)$	$(\sigma_\mu\ \sigma_\mu)$
ba da	sa li

x	x
$(\sigma_{\mu\mu})\sigma_\mu$	$\sigma_\mu(\sigma_{\mu\mu})$
faː tu	ɟe gaːn

In three- and four-syllable names, such as [majmuna] 'Maimouna' and [marijama] 'Mariama', the stress appears to fall three syllables in from the right edge. We will assume for now that the glide in the first syllable of [májmuna] makes the syllable heavy and that a foot is constructed around it to avoid a heavy-light trochee.

(2)

x	x
$(\sigma_{\mu\mu})\sigma_\mu\ \langle\sigma_\mu\rangle$	$\sigma_\mu\ (\sigma_\mu\ \sigma_\mu)\langle\sigma_\mu\rangle$
maj mu na	ma ri ja ma

If we were to avoid claiming extrametricality, the words in (2) would require building iambic feet right-to-left, but as we saw in (1), this cannot be the case, or else we would predict *[badá] and *[salí]. Thus, we know that extrametricality plays a role. As for directionality, it is possible that trochees are parsed left-to-right and initial syllables are extrametrical, which would still predict [maríjama]. If this is true, there would need to be a rule that foots the heavy syllable /maj/ in /majmuna/ first. Perhaps, words are first scanned left-to-right for heavy syllables to assign them stress, and if none are located, then extrametricality applies and regular footing resumes. This seems unlikely given the name [bíram] 'Biram'.

(3)

x	x
* $(\sigma_\mu\sigma_{\mu\mu})$	* $\sigma_\mu(\sigma_{\mu\mu})$
bi ram	bi ram

Under a left-to-right analysis, /biram/ would surface as [birám], while a right-to-left analysis would result in a light-heavy trochee. The problem here is the weight of the second syllable. If we consider the final /m/ to be non-moraic, both of these analises would be resolved. McLaughlin (2000) finds evidence for non-moraic codas in reduplication patterns. Reduplicants in Sereer must obey prosodic well-formedness constraints which enforce the profile of a bimoraic syllable. If the copied syllable is CVC, e.g. /kim/ 'song', the nasal is not considered and instead, the vowel lengthens to fit a CVV template: /kim-kim/ → [kiːgim] 'singer' (ignoring consonant mutation). In fact, there is reason to believe that word-final consonants are onsets of empty nuclei. This is beyond the scope of this study and we will go on to ignore words with final consonants. McLaughlin (2000) only provides examples of monosyllabic bases so syllable-final consonants that are not word-final are absent from her analysis. However, glides and nasals can be found word-medially as potential codas and these do seem to attract stress as in [madémba] 'Mademba'.

(4)

x	x
$\sigma_\mu\ (\sigma_{\mu\mu})\langle\sigma_\mu\rangle$	* $(\sigma_\mu\ \ \sigma_\mu)\ \langle\sigma_\mu\rangle$
ma dem ba	ma dem ba

Going back to the question of directionality, since we have proven extrametricality and the building of trochaic feet, under a right-to-left analysis, /roqaja/ 'Rokhaya' would incorrectly surface as *[roqája] and not as [róqaja].

(5)

x	x
$(\sigma_\mu\ \sigma_\mu)\ \langle\sigma_\mu\rangle$	* $\langle\sigma_\mu\rangle(\sigma_\mu\ \ \sigma_\mu)$
ro qa ja	ro qa ja

To summarize the findings up until this point, Sereer parses trochees from right-to-left and deems the rightmost syllable extrametrical whenever possible, avoiding building degenerate feet. We will illustrate with the names [ɲɟápali] 'Ndiapaly', [ramátulaj] 'Ramatoulaye', [ájsatu] 'Aissatou', and salímata 'Salimata'.

(6)

x
$(\sigma_{\mu}\ \sigma_{\mu})\ \langle\sigma_{\mu}\rangle$
ɲɟa pa li

x
$\sigma_{\mu}\ (\sigma_{\mu}\ \sigma_{\mu})\langle\sigma_{\mu\mu}\rangle$
ra ma tu laj

x
$(\sigma_{\mu\mu})\ \sigma_{\mu}\ \langle\sigma_{\mu}\rangle$
aj sa tu

x
$\sigma_{\mu}\ (\sigma_{\mu}\ \sigma_{\mu})\ \langle\sigma_{\mu}\rangle$
sa li ma ta

Notice that [sáli] is a truncated version of [salímata] and stress on [li] is not maintained for the hypocoristic. Rather, stress rules are reapplied once the original name is truncated. Importantly, hypocoristics must always be minimally the size of a foot, as with [sáli] from [salímata] and [fáːtu] from [fáːtumata]. It is also possible that the /mata/ ending is a diminutive suffix (or something similar) which falls inside the stress domain of the base name, and that stress rules are reapplied after the incorporation of the suffix. Whether something is added or removed, stress assignment enacts on the domain post-suffixation/truncation.

Looking at [fáːtumata] 'Fatoumata' and [ɟáːratulaj] 'Diaratoulaye', we see that stress is drawn to the long vowels and shifts one syllable to the left from where it would be predicted to be.

(7)

x
$(\sigma_{\mu\mu})\sigma_{\mu}\ \ \sigma_{\mu}\ \langle\sigma_{\mu}\rangle$
faː tu ma ta

x
$(\sigma_{\mu\mu})\sigma_{\mu}\ \ \sigma_{\mu}\langle\sigma_{\mu\mu}\rangle$
ɟaː ra tu laj

We pose that after the rightmost syllable is made extrametrical, the word is scanned right-to-left for the rightmost heavy syllable and builds a foot around it, leaving medial syllables unparsed. Because the longest nouns in the data are four syllables, we are unaware if the limit is four syllables in from the right until regular right-to-left footing ensues, or if a theoretical infinite number of unfooted light syllables can intervene until one heavy one is found. Also, since words longer than four syllables were not elicited (and may not exist), no evidence has been found for secondary stress.[4]

Finally, we have been staying away from verbs since most stems are either monosyllabic or compounds, and together with the morphology that surrounds them, tackling the dimensions of stress domains would be unreasonable to do in one study, so we will leave verbs for future research. One thing is certain however, stems are more prominent than their accompanying functional morphology. For example, the verb complex /a-gaʔ-a/ 3SG-*see*-PRES 'he/she sees' surfaces as [agáʔa]. The last syllable goes up in pitch when it is aligned with the right edge of a PPh (see §2.3) but it is still lower in intensity than the stem. Said in isolation, there is no final rise and the syllable is still less intense than the stem to its left.

When the same verb is in past tense, /a-gaʔ-ʔa/ 3SG-*see*-PST 'he/she saw', the final syllable is as loud as the stem. Even when said in isolation, the past tense morpheme /-ʔa/ has high amplitude. A parallel can be drawn with Spanish past tense markers carrying their independent stress such as [kóm-o] *eat*-1SG.PRES 'I eat' and [kom-í] *eat*-1SG.PST 'I ate' (Harris, 1992). Therefore, we posit that verb stems are lexical morphemes and are assigned stress, and some functional morphemes come with prespecified stress. Crucially, whether an H phrase tone is on the phrase-final syllable or not, intensity contours are maintained, remaining faithful to stress. Additionally, the tone always associates to the phrase-final syllable uninformed of where stress is, meaning that the two systems do not interact.

5.2 Intonational Structure

Now that we have established the mechanics of stress in Sereer and that it manipulates intensity as a signal, we will explore how the pitch dimension is controlled by intonational structure.

5.2.1 Constituency

Cross-linguistically, declaratives tend to be marked with an L% tone at the right edge of the utterance, and normally a H% tone at the left edge. This coincides with the fact that utterances begin higher and stronger, then due to declination and utterance-final weakening, the utterance ends lower and weaker (Pierrehumbert, 1980). Sereer is no exception to this contour. When single words such as /sali/ 'Sali' are said in isolation, utterance prosody applies to them and they take the surface form of [$^{\mathrm{H\%}}$sáli$^{\mathrm{L\%}}$] where the strong-weak stress pattern aligns with the HL intonational pattern. However, when the same word is aligned with the right edge of a medial PPh, which is typically associated with an H tone, the pitch contour changes to [sáli$^{\mathrm{H}}$]. To an untrained ear that listens out for pitch, this gives the illusion that the stress shifted to [salí], but in fact, stress still falls on the initial syllable, it is simply being marked with intensity.

[4]This is because a four syllable name like /salimata/ would need to be parsed as (sa)(lima)⟨ta⟩, with a degenerate foot around the initial syllable.

Short declarative sentences such as /ŋgor aɲaːma liɓ̥/ 'Ngor eats fish' do not need intonational subdivisions when said at a high speech rate, though they can be broken up if desired by the speaker. When the IP contains a single PPh, the output is [$^{H\%}$ŋgor ^{L}aɲaːma liɓ̥$^{L\%}$], with the highest peak on /ŋgor/ and a drop to a mid range immediately after on /a-/. From /a-/, the rest of the utterance has a steady downward slope in f0 until the L% target at the end where it drops to the lowest pitch. With phrase divisions, either the subject and verb complex, or the verb complex and object surface with no boundary between them as a single PPh. The sentence /majmuna agaʔa sali/ 'Maimouna sees Sali' can be structured as in Figure 7. We will remove feet, syllables, and moras

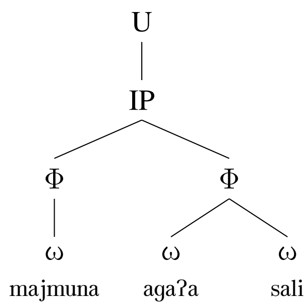

Figure 7: Prosodic structure of [(S)(VO)] parse: [$^{H\%}$maj^{L}muna^{H} ‖ ^{H}(ʔ)a^{L}gaʔa sali$^{L\%}$]

for space, but note that they are still in the structure. This prosodic parse renders a surface contour which places an audible boundary between the subject /majmuna/ and the predicate /agaʔa sali/. An H tone appears to mark the right edge of the first PPh, while an optional[5] H tone and an epenthetic glottal stop appear to mark the left edge of the second PPh. Boundary tones take place at either side of the IP.

Instead of the [(S)(VO)] parse, if this same sentence is uttered with an [(SV)(O)] parse, the structure and contour would change to that of Figure 8. Like before, this sentence has H tones and a pause marking the right edge of the fist PPh and the left edge of the second PPh, as well as boundary tones H% and L% at the left and right edge of the IP respectively. Interestingly, a glottal stop does not divide /majmuna/ and /agaʔa/ as we would expect if there were a boundary here. Instead, the

[5] H tones at the left edge of medial PPhs do not appear all the time. Usually, the H tone on the right edge of the preceding PPh and the pause are enough to mark the boundary, taking precedence over the optional H after the boundary.

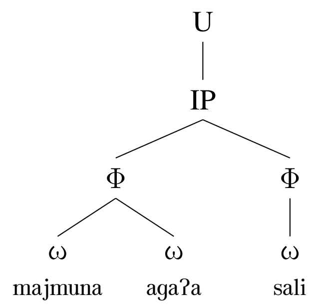

Figure 8: Prosodic structure of [(SV)(O)] parse: [$^{H\%}$majmuna agaʔa^{H}‖ ^{H}sali$^{L\%}$]

lack of a boundary leaves the two vowels in hiatus with the same phonetic length of an /aː/ except intensity and pitch drop for the second vowel, rather than staying flat like regular long vowels. We explain this drop on the pre-verbal element /a-/ '3sg' as a preparation for the prominence of the following verb stem /gaʔ/. If the second /a/ was the start of a new PPh, we would anticipate an H tone marking the start; not a slight drop in intensity and pitch. All of this indicates that the adjacent /a/ vowels are kept as separate segments divided by a PWd boundary yet not a PPh boundary.

5.2.2 H+H Tone Contact

When the initial PPh is built around three syllables, the H% tone on the left makes σ_1 high. The H tone on the right makes σ_3 high at around the same f0 or higher than σ_1. To keep these two tones distinct, σ_2 drops to an L at a relative mid-to-low range. However, when the phrase is two syllables long, this partition between high tones needs to be made elsewise. Figure 9 shows the pitch contour of an utterance with a bisyllabic initial PPh.

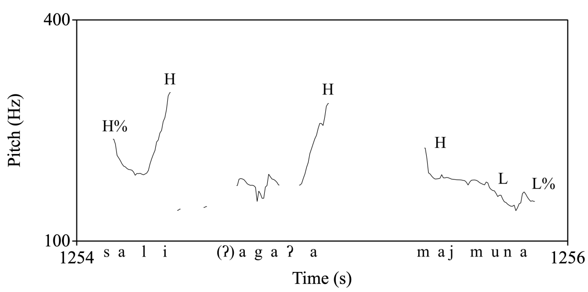

Figure 9: Annotated pitch graph for /sali agaʔa majmuna/ 'Sali sees Maimouna'

Note that the high-frequency frication of the initial [s] in /sali/ makes the initial syllable appear to start higher than it truly would exclusively considering f0. Comparatively, the H% tone still affects σ_1 but does not trigger the pitch to rise as much as with the following H, which skyrockets the pitch of σ_2 with respect to σ_1.[6] Though the first syllable of /sali/ is stressed, intonation reduces the hight of H% to make way for of the following H target, almost as if decreasing the prominence of σ_1 and increasing the prominence of σ_2 through pitch. But since stress is preoccupied with manipulating intensity, intonation is unconstrained to alter the pitch. Notice that an L phrase tone precedes the L% at the right edge of the utterance. We will come back to discuss this fact in the later part of this section.

5.2.3 Coordination and Pitch Accent

Take the sentence /sali fo majmuna aweɲaːmaː saːʄ̥/ 'Sali and Maimouna are eating sathie'[7] with two coordinated nouns in the initial PPh. At a relaxed pace, S, V, and O are each grouped into their own PPhs as indicated by the phrasal tones and caesuras.

H%	L				H		H	L				L%
|	|				|		|	|				|
sa	li	fo	maj	mu	na	‖	(ʔ)a	we	ɲaː	maː	‖	saːʄ̥

Figure 10: 'Sali and Maimouna are eating sathie'

This utterance is composed of three PPhs, the last of which contains only one syllable. In the case of monosyllabic final phrases, the L% tone defines the pitch of the syllable and takes precedence over the optional H tone that would normally be at the left edge of its phrase. Drawing our attention to the initial phrase which is built around six syllables, the H% tone on the left affects the first syllable of /sali/ and the H tone on the right affects the last syllable of /majmuna/. In this particular utterance of this sentence, every medial syllable after H% plateaus at a mid-low range until arriving at the H.

By contrast, it is also possible for this same sentence to have the contour in Figure 11. In this parse, the coordinator /fo/ 'and/with' is considerably higher in f0 than its neighbouring syllables. Though it is likely not stressed since it is

H%	L	H*	L		H			H		H	L	L%
|	|	|	|		|			|		|	|	|
sa	li	fo	maj	mu	na	‖	(ʔ)a	we	‖	ɲaː	maː	saːʄ̥

Figure 11: 'Sali and Maimouna are eating sathie'

not louder than its neighbouring unstressed syllables, it appears to mark the left edge of /fo majmuna/ 'and Maimouna.' At an exceptionally slow speech rate, a break is inserted between /sali/ and /fo/, and a smaller break is inserted between /fo/ and /majmuna/, suggesting that the coordinator attaches to the noun to its right. These breaks are not strong enough to be heard at a normal rate like PPh boundaries. Hence, /fo majmuna/ forms a PWd and can be assigned a pitch accent H* at its left edge to signal the coordination with the preceding PWd, both of which combine to form the larger PPh /sali fo majmuna/.

On top of that, at a slower speech rate, the 2pl progressive marker /awe/ has prosodic pauses on either side, takes an initial glottal stop, and is marked with an H tone at the right edge. It forms its own PPh rather than adjoining to the verb, which is interesting given that it constitutes verbal morphology. Since /awe/ can be singled out like this, we propose that it is not a clitic or a prefix. Though morphosyntactic analysis is beyond the scope of this study but for future directions, its prosodic standing shown in this paper may serve as evidence for its morphological proximity to the verb. The pitch accentuation of the unstressed item /fo/ suggests unstressed morphemes can be associated to prominent tones. If intonation was anchored by stresses, we would expect the H* to fall on the initial syllable of /majmuna/, not the coordinator.

5.2.4 Edge-Alignment Effects

Syllables aligned with left edges are known to be subject to strengthening and syllables aligned with right edges are known to weaken and sometimes erode. Initial syllables in Sereer tend to be louder, higher in f0, and marked by an epenthetic glottal stop when vowel-initial. Let us examine the final PPh of the following utterance:

The two elements /na wo/ 'for you', may form a PPh together depending on speech rate, as shown in 12, or they may join the preceding verb into a single PPh, as in the alternative parse [$^{H\%}$foːfiʔ^{H}

[6] The pitch on [sali] still descends before hitting the H at the edge and could constitute an H+L (falling) tone.

[7] Sathie is a dish similar to couscous but made with millet.

```
H%  H      H  L  H      H  L%
|   |      |  |  |      |  |
fo: fiʔ ‖ (ʔ)a faː ha ‖ na wo
```

Figure 12: ‘Water is good for you’

‖ ^{H}afaːha na^{L} wo$^{L\%}$]. Notably, the intensity of the preposition /na/ ‘for’ varies depending on the prosodic structure. When a prosodic boundary separates the verb complex from /na/, the average intensity of /na/ is 55.12dB. However, when the entire predicate forms a single PPh, /na/ is left in medial position, and its average intensity drops to 52.47dB.

In addition, the 2sg strong pronoun /wo/ ‘you’ consistently exhibits higher intensity than /na/ when both elements are part of a PPh with the preceding verb complex. When /na wo/ forms its own PPh, /na/ roughly matches /wo/ in intensity. Despite being final, /wo/ resists the typical utterance-final weakening and provokes a higher peak than a preceding phrase-medial /na/, and a roughly equal peak in loudness than a preceding phrase-initial /na/. In this type of instance, prosody interacts with stress, since the left-edgedness of /na/ strengthens the syllable, increasing intensity and interfering with the stress cue. Interestingly, both items have phonologically short vowels, but /wo/ is around twice the length (0.303ms) as that of /na/ (0.173ms). Though this can be due to final lengthening. This constancy in the intensity of /wo/, coupled with the variability of /na/’s intensity suggests that unstressed elements that lie on left edges of phrases may exhibit intensity levels that approach those of truly stressed elements. A parallel pattern can be observed with the preposition /a/ ‘to’ as in the phrase /a bada/ ‘to Bada’. When /a/ is not at the PPh edge, it has lower intensity than the stressed first syllable of [báda]. By contrast, when /a/ lines up with the left edge of the PPh, it displays the same intensity as the first syllable of [báda].

Thus, we conclude that strong pronouns like [wó] are stressed, and indicated by their consistently high loudness, particularly in utterance-final position. Other strong pronouns, such as /ten/ ‘3sg’ and /mi/ ‘1sg’ are also stressed and behave in the same way. Furthermore, prepositions like [na] and [a] are unstressed, with their stress-like loudness being contingent on their alignment with a phrase edge. In this construction we see that prosody, not just intonation, strengthens an unstressed syllable because of its left-edgedness, interfering with its relative prominence and blind to neighbouring lexical stresses fighting for prominence.

5.2.5 Utterance-Final L%

As mentioned previously, the final L% tends to bring down the pitch of preceding syllables near the right edge of the utterance. This has also been observed in Wolof, a related language (Rialland & Robert, 2001) In this subsection we will focus on these patterns. In the following sentence, the prosodic boundary between /ajsatu/ ‘Aissatou’ and /acija/ ‘gives’ is overtly marked. However, the boundary between /foːfiʔ/ ‘water’ and /bada/ ‘Bada’ is very small and likely just the space after the release of the glottal stop at the end of /foːfiʔ/ and the closure of the labial stop at the beginning of /bada/.

```
H%        H     H  L          H     H  L%
|         |     |  |          |     |  |
(ʔ)aj sa  tu |  a  cij  a fo: fiʔ ‖ ba da
```

Figure 13: ‘Aissatou gives water to Bada’

Nonetheless, the phrase break is still signaled by the H tones at the edge shared by the two phrases. This kind of construction where the final PPh is bisyllabic like /bada/, results in a sharp drop in pitch between the H tone at the left edge and the L tone at the end of the utterance, from 204Hz to 122Hz at its most drastic. This HL contour is comparable to that of the same bisyllabic word said in isolation. See Figure 4.

The second half of the same sentence can be said with an optional preposition /a/ ‘to’, which adds an extra syllable to the final phrase.

```
H%        H     H  L          H     H  L  L%
|         |     |  |          |     |  |  |
(ʔ)aj sa  tu ‖  a  cij  a fo: fiʔ ‖ a  ba da
```

Figure 14: ‘Aissatou gives water to Bada’

In this variation, the break between /foːfiʔ/ and /a/ is stronger and an epenthetic glottal stop is inserted to further signal the start of the final PPh. Here, the drop from H to L happens between /a/

and the first syllable of /bada/. In other words, the L tone at the end of the utterance shifts inwards when there are free syllables with no tones associated to them. See Figure 5.

In the trisyllabic case of /a bada/, the L tone at the right edge spreads inward (right-to-left) by one syllable before it comes into contact with the H tone at the other edge. The next question would be: how far does the L tone spread when the PPh is more than three syllables and the H tone is located further up the sentence? This could simply be due to an H to L drop, like in initial or medial phrases, except happening before the final L%. Either way, it contradicts an English-style trisyllabic final phrase, where the H tone at the left edge keeps the first two syllables relatively high, while the L% tone at the right edge only affects the final syllable. In other words, an English trisyllabic final phrase is HHL and a Sereer trisyllabic final phrase is HLL.

When this sentence is said at a high speech rate, the entire predicate can constitute one PPh.

H% H H L L%
(ʔ)aj sa tu ‖ a cij a foː fiʔ ba da

Figure 15: ‘Aissatou gives water to Bada’

An H% boundary tone appears at the right edge of /ajsatu/ and seemingly another at the left edge of /acija/, though the two vowels are not separated by a glottal stop as they would in slower speech. Thus, the boundary is weaker but still signaled by an f0 rise. The second PPh features the H target at the left edge on /a-/, a fall on /-ci-/, and gradual declination thereafter over the following three syllables /-ja foːfiʔ/. On the first syllable of /bada/, another drop occurs as if to foreshadow the final L%. This way, it becomes clearer that there is a second drop closer to the final L% that is unrelated to the drop caused by the L earlier in the phrase. See Figure 6.

If the L% tone is spreading in from the right, there are two possibilities: 1) either two syllables is the most a final L% is allowed to spread (the ultimate and penultimate syllables of the final phrase), or 2) the L tone can only affect up to the leftmost syllable of the final lexical item and no further. We conclude on the former given the final phrase [...^{H}a^{L}wara poko^{L}laʄ̥$^{L\%}$] ‘... kills a spider’, where /pokolaʄ̥/ ‘spider’ has the contour HLL. Though f0 is at a mid range after slowly declining following the L on /wa/, it is relatively H compared to the following two syllables, which drop to an L target on /la/. If the definite marker /-ne/ is added to the end of /pokolaʄ̥/, the output is [pokolaʄ̥^{L}ne$^{L\%}$] with a contour HHLL. Therefore, the extent of the L% spread does not appear to be informed by the edge of morphosyntactic domains, but rather counts two syllables in from the right. Altogether, these tone association rules pay attention to prosodic boundaries, number of syllables from a prosodic edge, and sometimes semantics with the case of coordination. Very little notice is given to the location of stress. As a result, the pitch and intensity cues rarely combine.

6 Conclusion

While the mismatch between pitch and intensity may create the illusion that Sereer is not a stress language, phonetic analysis and statistical tests seem to confirm that stress is present and cued by intensity. As for stress, Sereer parses left-to-right trochees with extrametrical syllables and prioritizes stress on heavy syllables. As for intonation, prosodic edges are marked by glottal stops, pauses, and boundary and phrase tones, with final L% tones drifting leftward by one syllable. Further, stress and intonation are largely independent from one another, and rely on different mechanisms that are uninformed of each other. In future research, more attention should be given to the stress patterns of verbs as it can give insight into morphology. With the information uncovered by this study, stress and intonation will be easier to disentangle, and the location of stress in newly elicited words will be more discernible.

Acknowledgments

The most enjoyable part of working on this paper was the opportunity to apply field methodology learned in class to elicit with our language consultant Thilor Ndiaye. We would like to thank her and both instructors Profs. Heather Goad and Martina Martinović who led us through this class. Thank you for all of your insight, thank you to Simon LiVolsi for helping us with Praat, and thank you to Thilor for giving us a chance to study her native language.

List of Sentences

1. ŋgor a-ɲaːm-a liɓ̥
   Ngor 3SG-eat-PRES fish
   ‘Ngor eats fish’
2. majmuna a-gaʔ-a sali
   Maimouna 3SG-see-PRES Sali
   ‘Maimouna sees Sali’
3. sali a-gaʔ-a majmuna
   Sali 3SG-see-PRES Maimouna
   ‘Sali sees Maimouna’
4. sali fo majmuna awe ɲaːm-aː saːʄ̥
   Sali and Maimouna 3SG.PROG eat-PRES sathie
   ‘Sali and Maimouna are eating sathie’
5. foːfiʔ a-faːh-a na wo
   water 3SG-good-PRES for 3SG.STR
   ‘Water is good for you’
6. ajsatu a-cij-a foːfiʔ a bada
   Aissatou 3SG-give-PRES water to Bada
   ‘Aissatou gives water to Bada’
7. biram a-war-a pokolaʄ̥
   Biram 3SG-kill-PRES spider
   ‘Biram kills a spider’

References

Becker, C. (1985). La représentation des sereer du nord-ouest dans les sources européennes (xve-xixe siècle). *Journal des africanistes*, *55*(1), 165–185. https://doi.org/https://doi.org/10.3406/jafr.1985.2093m

Ferraira, F. (1993). Creation of prosody during sentence production. *Phonological Review*, *100*(3), 233–253.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Basil Blackwell.

Harris, J. W. (1992). *Spanish stress: The extrametricality issue*. Indiana University Linguistics Club.

Hayes, B. P. (1980). *A metrical theory of stress rules* [Doctoral dissertation, Massachusetts Institute of Technology].

Hayes, B. P. (1989). Compensatory lengthening in moraic phonology. *Linguistic inquiry*, *20*(2), 253–306.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.

Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, *8*(2), 249–336.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, *32*(4), 451–454.

McCawley, J. D. (1978). Iv - what is a tone language? In V. A. Fromkin (Ed.), *Tone* (pp. 113–131). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-267350-4.50009-1

McLaughlin, F. (2000). Consonant mutation and reduplication in Seereer-Siin. *Phonology*, *17*(3), 333–363. http://www.jstor.org/stable/4420178

Merrill, J. T. (2023). *The Cangin languages: Phonological and morphological reconstruction and diachrony* (Vol. 21). Brill.

Nespor, M., & Vogel, I. (2007). Prosodic phonology: With a new foreword. In *Studies in generative grammar* (Vol. 28). Mouton de Gruyter.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* [Doctoral dissertation, Massachusetts Institute of Technology].

Prince, A. S. (1983). Relating to the grid. *Linguistic inquiry*, *14*(1), 19–100.

Rialland, A., & Robert, S. (2001). The intonational system of Wolof. *De Gruyter Mouton*, *39*(5), 893–939. https://doi.org/doi:10.1515/ling.2001.038

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology*, *3*, 371–405. https://doi.org/10.1017/S0952675700000695