Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

The interaction of lexical tone and phrase-level intonation in Limbum

Siri Gjersøe, Jude Nformi Awasom, Ludger Paschen

Leipzig University
{siri_moen.gjersoe, awasom_jude.nformi, ludger.paschen}@uni-leipzig.de

Abstract

This paper presents results of an acoustic study of tone in Limbum, a Grassfields language of Cameroon. Our main claim is that Limbum has a final low boundary tone (L%) that appears in phrase-final position in assertive sentences and wh-questions but not in polar questions. We present evidence that this boundary tone can have three different phonetic manifestations in Limbum: (i) lowering of low level tones, (ii) the emergence of falling contour tones, and (iii) breathy voice. As there is a strong correlation between the presence of L% and falling tones, we propose that these contour tones are in fact the result of a single tone combining with a low boundary tone. We thus challenge previous accounts assuming underlying contours that are simplified in pre-pausal position. The findings give rise to the question of how the contrast of stable level vs. alternating level/contour tones before a boundary is represented in the phonology.

Index Terms: tone language, boundary tone, contour tones, voice quality, Bantu, Limbum

1. Introduction

Limbum belongs to the Grassfields family (Bantu) within the Niger-Congo languages and is spoken by about 90,000 people in the Donga Mantung division of the North West region of Cameroon [1]. Limbum is an understudied tone language, and only a limited amount of descriptive works are available [2, 3, 1, 4]. To the authors' knowledge, no detailed study of the phonetics of Limbum tones has been carried out as of yet, and no attempts at describing the basic units of Limbum prosody (boundaries, intonation tunes) have been made before, either.

Limbum distinguishes three basic tone levels: H (high), M (mid), and L (low). In addition to these three level tones, Limbum has several contour tones, but the exact details as to the inventory and the distributional properties of the contour tones are still obscure, and descriptions in the literature are not unanimous. Thus, some sources identify five gliding tones (HL, HM, LM, ML, LL) [3, 1], while others identify only four (HM, ML, LL, LM) [2] or even two (HL, LH) contour tones [4]. Falling contour tones have been claimed to undergo neutralisation with the respective level tones in sentence-medial positions [1]. The claim that we will make here is that only a minority of the alleged contour tones are true lexical contour tones, and that falling contours arise due to the presence of a low boundary tone.

The paper is organised as follows: section 2 presents the setup and design of a production experiment that was conducted to investigate the acoustics of tone in Limbum. Section 3 discusses the results of the study for level tones, contour tones, and the sentence-final particle. Section 4 addresses some general issues regarding the interaction of tonal and prosodic phonology.

2. Methodology

2.1. Data

Data presented in this paper were collected from a 26-year old male native speaker of the Central (Linti) dialect of Limbum spoken in Nkambe. Recordings were made in a quiet, closed room at Leipzig University in Germany using a T-bone SC 440 supercardioid microphone. The informant read out aloud a list of test sentences containing a monosyllabic open-syllable target word which belonged to one of the eight tone classes as described by most available sources [3, 1] and as confirmed by our informant (see table 1; we excluded LM since the focus of this study was on level and falling tones only).[1] According to our informant, only one word with HM is attested in Limbum. For all other tones, we tested two different lexical items each, one ending in *-a* and one ending in a different vowel (see sec. 3.3 for motivations). The target words were tested in five different constructions (see ex. (1) below) to test for positional and pragmatic factors that we expected to affect presence and behaviour of boundary tones. Each item (= one sentence with a specific target word) was repeated three times, yielding a total of 65 items and 195 tokens.

Tone	Word 1	Gloss 1	Word 2	Gloss 2
H	*bá*	hill	*fú*	rat
M	*ntā*	market	*ŋgū*	wood
L	*shà*	kind of drink	*bì*	people
HM	*báā*	two	–	–
HL	*cwáà*	yellow bird	*kúù*	funnel
ML	*tāà*	father	*bīì*	co-wife
LL	*ràà*	bridge	*nkfùù*	bachelor

Table 1: *List of target words with different tones used as stimuli.*

In plain assertions (*Ass*, (1a)), the target word is in a sentence-final position, while in the control sentences (*Con*, (1b)), it is in a non-final position. We also tested different question types. Wh-questions in Limbum can occur with or without a sentence-final particle *a* (*Wh1* (1c) and *Wh2* (1d)). Polar questions (*Pol*, (1e)) obligatorily take the final particle *a*. Each sentence contained an initial H-toned word (*Tánkó,* á) and a sentence-medial M-toned word (yɛ̄) for pitch baseline reference.

(1) a. Tánkó àm yɛ̄ zhí **shà**
T. PST see POSS drink
'Tanko saw his drink.' (Assertion)

b. Tánkó àm yɛ̄ zhí **shà** fī
T. PST see POSS drink new
'Tanko saw his new drink.' (Control)

[1] Our transcription of segmental length reflects duration measurements in our data set and does not necessarily coincide with descriptions in other sources.

 doi: 10.21437/SpeechProsody.2016-23

c. á ndā àm yɛ̄ bá zhí **shà** a
FM1 who PST see FM2 POSS drink PRT
'Who saw his drink?' (Wh-Q 1)

d. á ndā àm yɛ̄ bá zhí **shà**
FM1 who PST see FM2 POSS drink
'Who saw his drink?' (Wh-Q 2)

e. Tánkó àm yɛ̄ zhí **shà** a
T. PST see POSS drink PRT
'Did Tanko see his drink?' (Polar question)

The recorded sentences were saved into .wav files and segmented using Audacity 2.1.0. The resulting files were transcribed with the help of the informant using Praat (v. 6.0.0.5) [5]. Afterwards, a Praat script was run to extract the F0 values of the first quarter, midpoint and third quarter of each annotated target and control vowel.

2.2. Hypotheses and quantitative analysis

The aim of the production experiment was to test whether the various tones and tunes described in the literature exist in Limbum, and how the tones interact with potential edge-bound prosodic phenomena (boundary tones). Another goal of the study was to test previous observations on contour tones alternating with level tones in sentence-medial position.

The results from the script were run in R studio (using R v. 3.2.2) for statistic analysis. We used Linear Models to test SentenceType (5 levels), Tone (7 levels), Target (3 levels: T(arget) word, P(article), and C(ontrol) word), and ToneType (2 levels: contour tone vs. level tone). The models were run on subsets of the data in order to obtain a useful comparison of SentenceType, Tone, and ToneType (contour tone vs. level tone). The dependent variables were *Midpoint* for level tones, and [First Quarter MINUS Third Quarter], henceforth *FirstQThirdQ*, for contour tones. Duration was also a dependent variable. First, all three level tones were tested on all sentence types with *Midpoint* as a dependent variable in order to test whether F0 values changed depending on the position of the target word (sentence-final vs. sentence-medial). Falling contour tones were checked for positional effects and also for possible influences of sentence type. The Null-hypothesis was that there is no F0 difference in target words across the different sentence types (F0 Midpoint for level tones and F0 contrast of contour tones FirstQThirdQ). In addition, we tested segmental length for the level tones and contour tones across the sentences types, and we measured the F0 Midpoint on the sentence-final particle in *Wh1* and *Pol*.

3. Results

3.1. Level tones

Table 2 shows the correlation between Midpoint and SentenceType for the three level tones L, M and H.[2] The F0 baseline (= Intercept) was the control sentence (*Con*) with the mean of H tone minus the grand mean. Its estimated coefficient was 129.7 Hz. The factor SentenceType explained a significant proportion of variance in Midpoint, rejecting the Null Hypothesis, $R2 = .697$, $F(6, 171) = 65.7$, $p < 0.001$. Target words in SentenceType *Ass* and *Wh2* had a significant decrease in Midpoint from the baseline (estimate = -9.8 Hz and -9.1 Hz). *Wh1* did not decrease significantly from baseline (-1.3 Hz), whereas *Pol* showed a (non-significant) increase in F0 compared to the baseline (2.2 Hz).

Figure 1 shows that in *Ass* and *Wh2*, the L tone (dark gray box) is considerably lower than in the other sentence types. The M tone (light gray box) and especially the H tone (black box) show weaker indications of lowering compared to L. This pattern was additionally tested in a separate Linear Model for each target tone, also with baseline of *Con*. SentenceType was a significant factor for Midpoint of L tones, $R2 = .516$, $p < .001$. The F0 value of *Ass* and *Wh2* decreased significantly in F0 (estimates = -11.9 Hz and -9.9 Hz, $p < .001$). For M on the other hand, SentenceType was non-significant, $R2 = .136$, $p = .14$. *Ass* and *Wh2* deviate non-significantly from the baseline (estimates = -7.3 Hz and -4.1 Hz, $p = .006$ and $p = .28$). For H, SentenceType was not significant, either, $R2 = .069$, $p = .52$. *Ass* and *Wh2* deviate non-significantly from the baseline (estimates = -6.6 Hz and -6.0 Hz, $p = .45$).

Factor	**Estimate**	**Std. Error**	**Pr(> $\|t\|$)**
(Intercept)	129.7	2.0	2e-16 ***
SenTypeAss	-9.8	2.8	0.0007 ***
SenTypeWh1	-1.3	3.6	0.7
SenTypeWh2	-9.1	2.8	0.001 **
SenTypePol	2.2	3.6	0.5

Table 2: *Results of the Linear Model on a subset of the target tones (H, M, L) testing the correlation between* Midpoint *(F0, Hz) and* SentenceType + Tone *(Intercept = Con).*

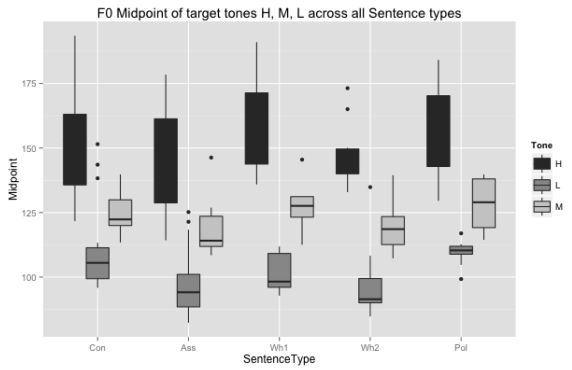

Figure 1: *Box plot of Midpoint (F0; y-axis) on H, M and L tone targets for each sentence type (x-axis).*

The results reveal that all level tones have lower F0 values in sentence-final position (*Ass* and *Wh2*) than in non-final position (*Wh1, Pol, Con*). However, differences were significant only for L, but not for M and H. Finally, polar questions show the highest F0 mean values of all level tones. Based on these results we propose that the lowering effect before a prosodic break in assertions and Wh-questions without particle is due to a final low boundary tone (L%) which is superimposed on the L lexical tone. Lowering is absent from *Pol* because L% is not licensed in polar questions. We tentatively propose that the non-significant lowering effects of H and M tone are due to downdrift. Finally, we attribute the higher Estimate for *Pol* to an overall higher pitch register in polar questions.

3.2. Contour tones

Limbum attests rising and falling contour tones. In this study, we examine the falling tones HL, HM, LM and LL (LL starts low and falls down to extra-low). Our data strongly support

[2] The results for each tone are excluded from the table for reasons of space. Results of separate tests on each tone are given instead.

the observation that low-falling contour tones are restricted to phrase-final positions, as all words that had a low-falling contour tone before a prosodic boundary were level-toned elsewhere. Figure 2 illustrates this dyadic behaviour of falling contour tones with pitch tracks from two tokens of the same word with an ML contour before a prosodic boundary (left) and a level M tone in phrase-medial position (right).

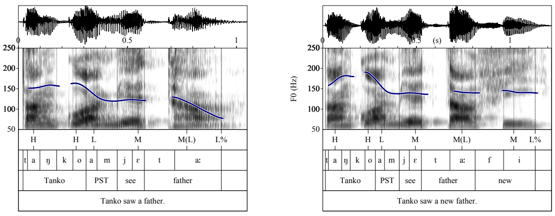

Figure 2: *Mid-falling contour tone on* taa *'father' before a pause (left) and a level mid tone on the same item in phrase-medial position (right).*

The F0 contrast between the first and third quarter of the target words was tested on a subset of falling tones for all sentence types. The results are given in table 3. The baseline was *Con* (estimate = 6.2 Hz). The factor SentenceType explained a significant proportion of variance in the F0 contrast of falling tones (FirstQThirdQ), R2 = .677, F(7, 107) = 32.1, $p < .001$. FirstQThirdQ values were significantly higher in *Ass* and *Wh2* where the target word was sentence-final. The contrast was non-significant in *Pol* and *Wh1* where the target word was followed by the question particle. The baseline estimate (*Con*) was lower than *Ass* and *Wh2*.

Factor	Estimate	Std. Error	Pr(> $\|t\|$)
(Intercept)	6.2	1.5	9.85e-05 ***
SenTypeAss	15.8	2.0	1.67e-11 ***
SenTypeWh1	-1.7	2.5	0.4
SenTypeWh2	17.7	2.1	1.67e-13 ***
SenTypePol	-0.5	2.3	0.8

Table 3: *Results of Linear Model of a subset with falling tones HL, HM, LM and LL. The model tested the correlation between* FirstQThirdQ *and* SentenceType + Tone *(Intercept = Con).*

The box plot in figure 3 illustrates the pitch slopes depending on different sentence types in more detail. The y-axis shows the difference in F0 at the 1/4 and 3/4 points of a target vowel and the x-axis shows the tested sentence types. As with the level tones, a considerable amount of variation for the different falling tones across the sentence types can be observed. However, one can clearly see that there is a distinct fall of up to 40 Hz in *Ass* and *Wh2* while the difference between the two measured points is much smaller in *Pol*, *Wh1* and *Con*, amounting to no more than 15 Hz. The main conclusion to draw from these figures is that the "falling" tones actually have a falling contour only before a prosodic break, but they do not have a falling slope when they occur in a phrase-medial position (*Con*, *Pol* and *Wh1*).

The data discussed in this section are ambiguous between the sandhi account, according to which phrase-final HM, HL, ML and LL tones are derived from underlying complex tones which lose their second (lower) component in all non-final positions [1], and an analysis that ascribes the final lowering pattern to a final low boundary tone. Since there is already independent evidence for L% affecting the L level tone before

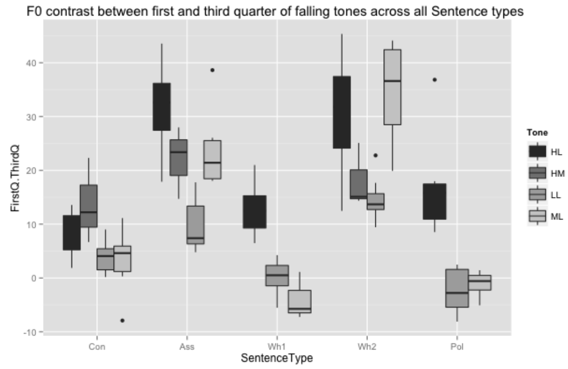

Figure 3: *Box plot showing F0 difference between First Quarter and Third Quarter of falling tones for all sentence types.*

a prosodic break, it appears reasonable to pursue a unified account for lowering of level tones and distributional restrictions on falling contours.

If we assume L% to be responsible for final falling contours, we still need to be able to distinguish between tones with and without a phrase-final falling contour, i.e. between stable level tones and level tones alternating with contour tones phrase-finally. One important factor for this seems to be segmental length. We found a positive correlation between Duration and the interaction of factors SentenceType and ToneType (contour tones vs. level tones), R2 = .369, F(9, 316) = 20.5, $p < .001$ (Baseline = *Con* with the mean of level tone minus the grand mean, estimate = 0.11 s(econds)). Level tones had non-significant values across all sentence types. Values for contour tones, on the other hand, were significant for *Ass* and *Wh2* (estimate = 0.03 s, $p < .001$, and estimate = 0.04 s, $p < .001$), and non-significant for *Con* (estimate = 0.009 s, p = .2). This means that while vowel duration is stable for plain level tones across different sentence types, it changes considerably with contour tones: vowels with a contour tone are longer in sentence-final position and shorter in sentence-medial position, where the tone also changes to level.

In sum, our results indicate that duration together with sentence position are two constraining factors that license a contour tone. We suggest that lexical items which display a falling contour tone phrase-finally should be modelled as having one specified and one unspecified tonal root node associated to the TBU (the syllable) $T_1(T_2)$ [6]. The boundary tone can dock onto $T_1(T_2)$ toned syllables to create surface contour tones; simple T tones, however, lack the second tonal root node and therefore cannot have an additional low tone link to the TBU and thus never become contour tones. Note that our analysis predicts that mid-falling tones could not be simply derived from low boundary tones, and indeed, H(M) appears to be an exceptional configuration restricted to a small set of lexical items (see section 2.1). In sentence types that lack L% (*Pol*), there is no tone to induce a low-falling contour. The reasons for the unexpected behaviour of (*Wh1*) lie in the presence of the final particle *a*, as will be discussed in the next section.

3.3. Questions and the particle *a*

The tone of the sentence-final particle *a*, which occurs in wh-questions and polar questions (see ex. (1)), was tested to check if it was also subject to L%. We tested the correlation between F0 Midpoint and SentenceTypes for the different targets (two levels: T(arget) word and P(article)). The baseline was SentenceType *Wh1* with the mean of target tone minus the grand mean (estimate = 143 Hz). We included only the level tones

(H, M, L) of the Word 2 column in table 1 because the particle *a* merges with the last vowel of *a*-final words which makes it impossible to test the tone on the target word against that on the particle. The results showed significant F0 differences for the interaction between the factors SentenceType and Target, R2 = .223, $F(5, 386) = 22.21$, $p < .001$. In *Pol*, T and P had about the same F0 midpoint (estimate = -15.0 Hz and -18.8 Hz). For *Wh1* on the other hand, the particle had a much lower F0 than the target word: T: estimate = -18.7 Hz vs. P: -51.6 Hz ($p < .001$). This means that the target word and the particle in polar questions have about the same F0 value while P in *Wh1* was significantly lower in F0 compared to T. In other words, the particle was considerably lower than the target word in SentenceType *Wh1* but not in *Pol*. The results are illustrated in figure 4. Similar results were obtained for the particle following a contour tone.

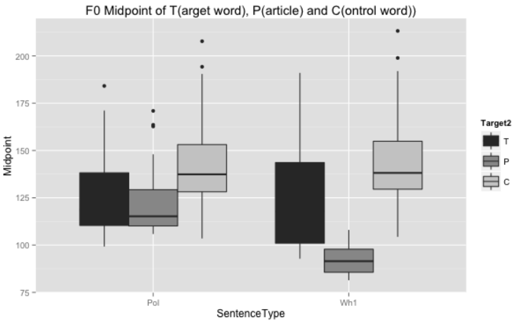

Figure 4: *Box plot of F0 midpoint of the p(article) compared with t(arget) and c(ontrol) words of level tones H, M and L in polar question and Wh1-question.*

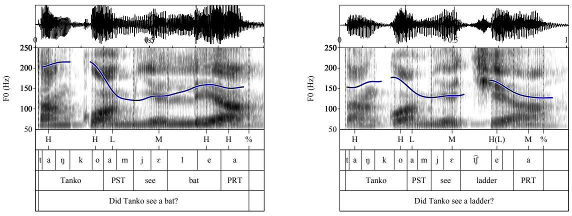

Figure 5: *Right phrase edge of polar question with H (left) and H(L) (right) toned item preceding the question particle.*

The results indicate that the particle *a* is affected by L% in wh-questions but not in polar questions. Our interpretation is that the particle *a* is tonally underspecified. When there is a boundary tone, the particle receives its tonal specification from the boundary tone (and becomes L-toned, see fig. 4); if the preceding item in the phrase is low-toned (L or L(L)), voice quality on the particle changes to breathy voice (see section 3.4). If there is a lexical level tone but no boundary tone, the lexical tone spreads to the particle. If there is a lexical $T_1(T_2)$ tone and no boundary tone, the lexical tone cannot spread to the particle due to the intervening empty root node that blocks spreading and M is inserted as a default tone (cf. fig. 5).

3.4. Breathy voice

A peculiar phonetic manifestation of L% in Limbum is breathy voice. Breathy voice usually occurs on final L(L) syllables and is only rarely found in non-low toned syllables, suggesting that it is a phonetic strategy to mark a low final boundary tone when a speaker is already at the lower edge of his voice range in a low-toned environment. Co-occurrences of low tones and breathy voice have been observed for a number of languages [7]. Figure 6 illustrates breathy and modal phonation before L% in L(L) and H(L) toned words.

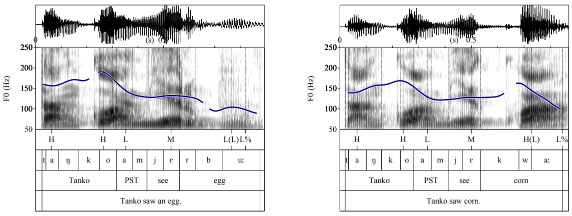

Figure 6: *Breathy (left) and modal (right) voice in different tonal environments: L(L) L% (left) and H(L) L% (right). The breathy vowel is characterised by a noisy spectrum and overall lower energy.*

4. Discussion

Despite the growing body of descriptive work on intonation systems in the languages of the world, the interplay of intonation and lexical tone is still understudied, and many puzzles have yet to be solved; this is especially true of tone languages with moderately or highly complex tone systems. In previous research, primarily two types of questions have been addressed: first, whether intonation tunes exist in a particular language at all [8], and second, how tonal contrasts are preserved in different phrasal environments [9] and how conflicting tonal specifications (e.g. lexical L and prosodic H*) are resolved. Curiously, in the case of Limbum, the question is not whether tonal contrasts are preserved in different prosodic contexts, but in which environments tonal contrasts become evident on the surface. Furthermore, our data suggest that Limbum follows a general trend to preserve prosodic tunes, however not by overriding or readjusting lexical material [10], but, in the case of L% preceded by L(L), by a special phonation. The fact that L% affects only L tones can be interpreted as a strategy for L% to preserve lexical tone contrasts. We hope that this observation sheds new light on the ongoing discussion of low boundary tones in Bantu languages ([11, 12, 13, 14], among many others).

The present paper also argued for a new analysis of the level/contour tone alternation. Unstable tones are a widespread phenomenon in Grassfield languages and are commonly analysed as products of tone sandhi rules [1]. The alternative account that we put forward does not rely on such stipulations, but instead considers the attested patterns as effects of lexical tone interacting with general prosodic processes.

5. Conclusion

This paper presented the results of an acoustic study of tones in Limbum, a Grassfields language of Cameroon. We assume a low boundary tone L% in all tested sentence types except for polar questions. L% affects L-toned words, words with two tonal nodes $T_1(T_2)$, and toneless particles, and induces breathy voice in phrase-final L(L) words. We argued that low-falling contour tones result from a complex interplay of lexical tonally underspecified tones and prosodically introduced tones. We thus offered an alternative account to what is sometimes referred to as tonal sandhi in some Bantu languages.

6. Acknowledgements

This research project was kindly supported by the Leipzig graduate school IGRA. We are grateful to Lukas Urmoneit for assistance in the recording sessions and Matías Guzmán Naranjo for consultation on the statistical parts of the study.

7. References

[1] M. A. E. Fransen, *A grammar of Limbum: A Grassfields Bantu language spoken in the North-West province of Cameroon*. [S.l.: s.n.], 1995.

[2] L. Fiore, *A phonology of Limbum*. SIL Cameroon, 1987.

[3] V. Bradley, *Description of surface tone in the Limbum verb phrase*. SIL Cameroon, 1994.

[4] S. T. Nforgwei, "A study of the phonological, morphological and syntactic processes in the standardisation of Limbum," Ph.D. dissertation, University of Yaounde I, 2004.

[5] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]," 2015. [Online]. Available: www.praat.org

[6] D. Odden, "Adjacency parameters in phonology," *Language*, vol. 70, pp. 289–330, 1994.

[7] J. Gao and P. Hallé, "The role of voice quality in Shanghai tone perception," *Proceedings of ICPhS 18*, 2015.

[8] S.-h. Peng, M. K. Chan, C.-y. Tseng, T. Huang, O. J. Lee, and M. E. Beckmann, "Towards a pan-Mandarin system for prosodic transcription," in *Prosodic typology*, S.-A. Jun, Ed. Oxford: Univ. Press, 2005, pp. 230–270.

[9] C. Gussenhoven, "Complex intonation near the tonal isogloss in the Netherlands," in *Prosodic typology II*, ser. Oxford linguistics, S.-A. Jun, Ed. Oxford: Univ. Press, 2014, pp. 324–364.

[10] W. Y. P. Wong, M. K. Chan, and M. E. Beckmann, "An autosegmental-metrical analysis and prosodic annotation conventions for Cantonese," in *Prosodic typology*, S.-A. Jun, Ed. Oxford: Univ. Press, 2005, pp. 271–300.

[11] L. Hyman. and R. Byarushengo, "A model of Haya tonology," in *Autosegmental Studies in Bantu Tone*, G. N. Clements and J. Goldsmith, Eds. Dordrecht, Foris, 1984, pp. 53–103.

[12] J.-M. Beltzung, A. Rialland, and E. Aborobongui, "Les relatives possessives en Embosi (C25)," in *Papers from the Workshop on Bantu relative clauses*, L. Downing, A. Rialland, J.-M. Beltzung, S. Manus, C. Patin, and K. Riedel, Eds. ZAS Papers in Linguistics, 2010, pp. 7–31.

[13] L. Hyman, "Do tones have features?" *UC Berkeley Phonology Lab Annual Report*, 2010.

[14] G. M. E. Aborobongui, F. Hamlaoui, and A. Rialland, "Syntactic and prosodic aspects of left and right dislocation in Embosi (Bantu C25)," in *ZASPiL 57. Proceedings of the Workshop BantuSynPhonIS : Preverbal Domain(s)*. ZAS Papers in Linguistics, 2014, pp. 26–48.