How intonations interact with tones in Embosi (Bantu C25), a two-tone language without downdrift

Annie Rialland and Martial Embanga Aborobongui
Sketch of tone system

- Two tones: H, L
- One tone per mora
- No tone spreading, floating tones, downstep
- Tonal rules are local and mainly triggered by vowel elision and glide formation.
- Generally, when a L tone and a H tone compete for a single mora, the H tone wins but this is not always the case, as exemplified in 2b.

(2) a. morábve < moro ábve “somebody fell”
   b. morakósi < moro ákósi “somebody hurts himself”
- Metotony:
  Thus, in present and future, verbs exhibit a final H tone when they are followed by a complement, or a L tone when they are not followed by a complement

(1) a. Monomoraic roots:
   H a-kó “cl9. forest”
   ω-tá “cl3. gun”
   L o-mbo “cl3. now”
   o-dzɛ “cl3. mockery”

b. Bimoraic roots (with a long vowel)
   HH i-báá “cl5. knife”
   LL i-baa “cl5. man”
   HL i-báa “cl5. marriage”
   LH i-báá “cl8. walls”

c. Bimoraic roots (with two syllables)
   HH ω-mbóndɔ “cl3. leg”
   LL i-mbamba “cl5. frog”
   HL o-lómi “cl1. husband”
   LH o-kondó “cl3. tail”
No overall downtrend

- No downdrift, downstep, declination
- H, L levels don’t change until final lowering
No Downdrift: Evidence

- What are slopes??

Table 1: Mean values with standard deviations of Hn and H(n+1) values for Speaker MEA and Speaker GNK

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Mean of Hn</th>
<th>Mean of H(n+1)</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEA</td>
<td>215Hz</td>
<td>216Hz</td>
<td>18Hz</td>
</tr>
<tr>
<td>GNK</td>
<td>189Hz</td>
<td>185Hz</td>
<td>4Hz</td>
</tr>
</tbody>
</table>

For both speakers, despite the dispersion of the pitch realization, which is mainly due to variations in register spans (see 3.3.), the mean values of Hn and H(n+1) are very close. For speaker GNK, there is a slight decaying tendency (around 4 Hz). But this slight decay is very small compared to the steps of a downdrift. For example, in Yoruba, a language with downdrift, the difference between a first H and a second H separated by a L tone is roughly 30Hz for speakers with similar pitch range (Laniran 1992). Based on this data, we can conclude that there is no downdrift in these utterances.

We can also remark that the standard deviation is more important for Hn, which is often the first H in a sentence than it is for H(n+1) with speaker MEA. This can be explained by the fact that the first H tone tends to be carried by a grammatical word and that grammatical words tend to display more variation than verbs or nouns, as they are occasionally reduced.

We next consider an important source of variation in tone realization: variations in the register span.
Final Lowering

- Rialland models as L% as an extra-low tone (at bottom of pitch range) that pulls the level of tones down, particularly Ls.
- But they also have temporal scope…
- What is temporal scope?
Final Lowering

- To me, it looks like the scope a contiguous sequence of (H) tones or sequence of (H)(L).
- L preceding the (H) or (H)(L) is not affected.
Models of final lowering

• Ignore the exact temporal scope (since it is unclear)
  • Assume 3-5 moras

(1) L% Gesture model

• Low-stiffness L% gesture aligned so reach its target in the last mora
• Target blends with Tone targets of overlapping Tone gestures
  Blending strength of L% with L > H (why?)
• Activation function of L% gesture (like pi-gesture)
• Possible Problem: Effect of initial conditions (H vs. L) when L% gesture kicks in.
  • Perhaps always triggered by a L, which fits with data
Models of final lowering

(2) Extension of Mambila Linear model

- Attraction of tone target to super-low L% located at end of IP
- Attraction is inversely proportional to:
  - F0 distance of tone target to L%
  - Temporal distance of tone target to L%

- One problem with both models
  - The final lowering on the surface is a smoothly falling f0 over the last X moras, beginning at or near the level of H.
  - This isn't explicit in either model.
  - This would account for why the final lowering scope is bounded by a L.
H% : Juxtaposed declaratives

- Extra High H% near end of IP.
- Attracted to H nearest the end of IP.
- Effect of H% seems pretty localized to one syllable.
- This would seem to require a nonlinear dynamics attracting the H% from the boundary to the first H back from the boundary.

---

In a given utterance, H and L tones stay essentially at the same level. Two reference lines can be posited: one for the H tones and another for the L tones. The L% tone is realized lower, toward the bottom of the speaker's pitch range. This organization can be schematized as follows in Table 2:

| H reference line | H_______H________H________ |
| L reference line | _____L________L_________L__ |
| L% bottom of the pitch range | _________________________L% |

As we examine juxtaposed assertive sentences (in section 4.6) and yes-no questions (in section 5), we will show that there are also extra-high intonational tones, which will provide a type of symmetry to the whole system, with the span of the tonal realization in the middle.

4.6 Sequences of juxtaposed assertive sentences:

Juxtaposed assertive sentences are very frequent in Embosi narratives. Each of them forms an intonational phrase, ending either with a L% or a H%. Starting with examples extracted from the beginning of an interview conducted by the second author, we introduce the H% and show some of its properties. At the end of a sentence, H% plays a similar role as the continuation rise in languages such as English and can be found in a similar context. In Embosi, its specificity is to be attracted by the last H of the intonational phrase. Consequently, it is realized on the last mora, which bears a H tone in example 8 and in Figure 6. The result is an extra-high tone at the end of the sentence and a rising contour that sounds like a continuation rise. In example 9 and in Figure 7, it is realized as an extra-high tone on the penultimate syllable, which bears the last H tone. Thus, the final contour is falling.

(8) [ngá bá mwánájooko b ap ñ réalisájéédijé]
ngá bá mwána áj ɛɛ oko b ap ñ réalisájéé di jé
1sgPRO here cl1.child cl1.come.PAST cl15.take.Hm. cl.9.stories cl9.REL. be.PRES cl9.DEM.
(litt.) “To me, here, the child came to collect stories which are these (ones)...”
‘To me, here, the child came to collect stories from here.’

(9) [ádzaábísímóondzési]
ádzaá bísí mó (m)ondzési
= it was us during childhood
= “It was during our childhood...”
Yes/No Q: H L%

L% is found at end of the utterance.

H occurs on the last H tone if the final tone is a L tone.

Otherwise, the H% of the HL% is pushed forward towards the preceding H tone, extending the contour on a larger domain.
Yes/No Q: H L%

- Modeling notes:
- The L% here is local. Doesn’t spread leftward.
- The effect of H seems local too. As with the juxtaposed H% there needs to be some nonlinear selection, as in the juxtaposed case.
- F0 slopes of assertions and yes-no Q are distinct?
- Theories of boundary tones?
Reference Lines

Table 3: Pitch references lines of H and L tones and L% and H% boundary tones

<table>
<thead>
<tr>
<th></th>
<th>H%</th>
<th>H% reference line</th>
<th>L reference line</th>
<th>L% bottom of the pitch range</th>
</tr>
</thead>
<tbody>
<tr>
<td>H%</td>
<td>____________________</td>
<td>H%</td>
<td>H_____H_____H_____</td>
<td>L%</td>
</tr>
<tr>
<td>H reference line</td>
<td></td>
<td>H_____H_____H_____</td>
<td>__<strong>L____L____L</strong></td>
<td></td>
</tr>
<tr>
<td>L reference line</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L% bottom of the pitch range</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- We can think of these as the Target values whose dynamics we modeled in the Mambila simulation.
Focus and wh-questions

- No intonation marking of focus or wh questions,
Emphasis: register expansion

- Raises H. Does it also lower L?
- “Emphasis (or backgrounding) modulates locally the informative weight.” In a continuous fashion.
Complex Sentences

- In corpus of isolated words and sentences, which include 45 complex sentences, there were no pauses or any other prosodic events between subordinate clauses and main clauses.

- Metotony in verbs: H->L in clause-final position.

(18) [ojúluláalámbatsínadílandzaa]  
(m)o-júlu láalámba tsína adí la N-jaa  
cl1-woman cl1.PRES-cook because cl1.PRES-be.Hm with cl9-hunger  
‘The woman cooks because she is hungry.’

L clause-final  H before complement
Left and right dislocations

**LEFT**

(19) [ojúluláαβɛлópɛwamúα L%/ɔkwáí L%] (Embanga Aborobongui & al. 2014)

(m)ο-γούλου 1-αβɛλά α-πέ 1-αβɛλά 1-γούλου 1-αβɛλά
1-woman 1.PRES.can.Hm cl15-give 1a.PRO 3sg.PRO 3-machete

‘The woman can give him it, the machete.’

**RIGHT**

Pause with no L% lengthening?

(20) [báso póo /báabáαβɛlότονάνο]

bási ó pøo 1-báa 1-báaβɛlα o-tønα 1-ø.
cl2.woman at cl7.village cl2.PRO cl2.can.PRES.Hm cl15-refuse.Hm 2sgPRO

‘The women, at the village, they can refuse you.’

(21) [Dislocated element [core clause]]IP