PHONOLOGICAL INTERPRETATION OF F_0 VARIATIONS IN A BANTU LANGUAGE: KINYARWANDA.

Thierry Chambon

Institut de Phonétique, Aix-en-Provence, France.

ABSTRACT

In order to interpret what is transcribed in the various studies of tone in Bantu languages by finding some points of reference in the acoustic signal, and to investigate the interaction between lexical tonal patterns and phrase intonational patterns, a recorded corpus of Kinyarwanda has been processed with a signal edition software, with which we can better interpret the relationship between the patterns, using a sophisticated method to measure and manipulate their respective features.

Pitch variations in so-called tone languages have generally been represented by associating segmental or autosegmental tonal features with particular syllables or moras. By modeling F_0 curves of Kinyarwanda, we examine which acoustic variations, among the various representations of the melody of this language proposed by linguists, are considered pertinent. Then, we can suggest a method to reconstitute the structure of its prosodic constituents and to sketch a production model of F_0 variations.

1. PROCEDURE

1.1 Corpus

We selected a set of nominal words with dissyllabic radicals representing all of the syllabic and tonal patterns found in the previous literature on this topic. Their morphology is structured as follows:
i-bi-múga (the disabled persons)
a-ba-gaanga (the doctors)
pre-prefix + noun class prefix + radical
These words were combined into utterances following the pattern :
N1 + beéretse + N2 + N3
(N1 showed N2 to N3).

1.2 Informants and Recordings.

The language of informants was representative of the Kinyarwanda spoken in the southern part of Rwanda. These informants were recorded in an anechoic chamber with a professional recorder and microphone. Utterances were typed on cards following the current Kinyarwanda spelling, namely without any vocalic quantity or accent marks, and presented to informants in a random order, except for a set of utterances composed only of radicals bearing no lexical H tone, to which signs of pauses (#) were added. The informants were told to group the words according to these marks (which sounded very natural to them, since these pauses never cut any tonal unit – see 2 –, but rather organized the sentence into intonative units as they had naturally spoken them before).

1.3 Operating and Modelling of Data.

These recordings were transferred from their analogic support to a digital one, namely the mass-memory of a MassComp 16 byte minicomputer, with a 10 kHz sampling rate. Using VES, signal edition software designed by R. Espesser & O. Balfourier, we edited oscillograms and sonagrams representing each sentence, segmented and labelled them with syllable markers, and validated the segmentation by listening to each labelled segment. VES can also be used to construct F_0 curves. We edited each curve into MOMEL (melody modeling software) and drew a superimposed F_0 curve by interpolating a quadratic spline function between target points located on F_0 peaks and valleys. This procedure, described in [2] enabled us to separate from the raw curve the (prosodic) intentional variations and the non-intentional variations, which result from articulation and coarticulation constraints – cf.[1]. Electromyographic studies indicated that in two different words showing the same stress structure but different segments (e.g. French [papa] and [mamã]), the vocal fold tensor muscles are preprogrammed in the same manner. Thus, the curve represents the phonologically relevant variations and gives an indication of the vocal fold control, where the inflection points correspond to the maximum muscle activity. To validate these abstract F_0 curves, utterances were recoded and played back to the informants. They could not hear any difference between the recoded and the original signals.

2. EXTRACTING THE PROSODIC CONSTITUENTS

We labelled these F_0 targets, assigning H (High) to F_0 peaks, L (Low) to valleys, and D to downstepped peaks – cf.[2]. These labels should not be interpreted like the ones in traditional tonological descriptions as they are not tone features associated with any syllable or segment, but relative pitches that describe contours of prosodic constituents disregarding their segmental structure. According to [3], prosodic structure can be represented using the following constituents: Syllable, Tonal Unit (T.U.) and Intonative Unit (I.U.). In accordance with the previous studies on Bantu languages, we will treat the syllable structure as: V, CV (one mora) or CVV (two moras): u-mu-saa-mbi (an old mat).

2.1 Tonal Unit

We assume for now that the T.U. is equivalent to the prosodic word, namely the unit that bears a lexical tone pattern (e.g. the radical and its extensions for short words). Then, the utterances in our corpus contain four T.U.: $[\text{abagabo}]_1$ $[\text{beéretse}]_2$ $[\text{abagoré}]_3$ $[\text{ibimúga}]_4$. We first grouped the words according to the seven possible tonal and moraic patterns found in previous descriptions – [4], [6] and others. By comparing them with the acoustic realizations, we tried to find a correspondence between the various transcriptions and discover the acoustic feature that was considered pertinent in each study (see 2.3).

2.2 Intonative Unit

Pauses are the most obvious clue in the identification of I.U., as the speakers group the utterance in one, two or three T.U. clusters. Moreover, in Kinyarwanda, a word final vowel is deleted when the following word begins with a vowel:
beéretse + abagabo = beéretsʼabagabo
(they showed the men)
beéretse + abagaanga + ibimúga = beéretsʼabagaangʼibimúga
(they showed the doctors the disabled people).
We assumed that the lack of application of this rule between two T.U. implies that there is an I.U. boundary. In order to verify this and to find the possible I.U. patterns in our corpus, we proceed as follows:
1) determine the pattern of I.U. bounded by pauses;
2) check that this pattern applies to any unit not bounded by a final vowel deletion (or any relevant segmental rule for other languages);
3) use this template to determine possible I.U. boundaries in the contexts in which the final vowel deletion can not apply: abageenzi + beéretse... (the travellers showed...).
To segment the I.U. when there is no pause, we need to know both its tonal template and its interaction with the U.T. templates. T.U. with the [LL] pattern (e.g. abagabo) never bear any F_0 peaks in the final position or when the final vowel deletion rule applies (fig.1). But an F_0 peak always appears on the last syllable before a pause or the lack of application of the final vowel deletion rule (fig.2). From this, we induced that there are two patterns for I.U. in this corpus: final [LL], and continuative [LH].

2.3 Interaction between Levels of Representations

Any T.U. in an utterance is necessarily supported by an I.U. We observed the realization of each word in our corpus in the context of both types of I.U. The final contour is dependent on three parameters: the underlying (lexical) pattern of the radical, the I.U. pattern, and some possible interactional and contextual rules (downstep). Our observations led us to extract only three possible T.U. patterns for this corpus: [LL], [HL], [LH]. The most interesting interactions observed are the ones involving an H in I.U. and T.U. simultaneously:

I.U.	[L H]	[L H]
T.U.	[LH]	[HL]
linearization	[LHH]	[LHLH]
Downstep1	[LHD]	[LHLD]
Downstep2 (optional)		[L H D]

This downstep is generally linked to the rising of the intermediate L. The regression of this F_0 valley in relation to the distance between the two surrounding peaks (d) gives a negative correlation (r= -0.609, a=0.01) and its value tends to meet the one of the second peak (to disappear) when (d) tends towards 150 ms. This explains the optional step of the rule and means that downstep seems to be triggered by a physiological rather than phonological constraint.
Without any H-interaction, we have the following linearizations:

I.U.	[L L]	[L L]	[L L]	[L H]
T.U.	[LL]	[LH]	[HL]	[LL]
lin.	[L L]	[LHL]	[LHL]	[LLH]

3. IDENTIFYING THE LEXICAL TONAL PATTERNS.

The difference between the [LH] and [HL] tonal patterns, which are both produced as [LHL] on isolated words, is difficult to evaluate because, according to our own and all previous research, there is no minimal pair illustrating such an opposition in Kinyarwanda. We grouped the words according to: a) the pattern [HL] or [LH] of the radical; b) the glissando on the first syllable of the radical (rising or falling); and we analysed in both cases the variance of the distance from the F_0 peak to the beginning of the word and of the radical. Results are significant in both cases. The difference based on the first syllable glissando seems to be the most attested (F=51, p=0.0001), and this explains why this has been the relevant clue in the previous studies. But, if this opposition seems to be the marked one for perception, which is predictable since glissando on vowels is the relevant feature for tone perception (cf.[7]), it could be just a consequence of the word pattern variation which may be more relevant for a production model. To find out which is the relevant domain for this opposition, we plan to collect a corpus of words beginning with the same class prefixes (a-ba vs i-bi) and possibly with the same radical initial syllable, in order to avoid the variations of the intrinsic and cointrinsic durations of segments. If we compare series of words grouped according to prefix duration and bearing the same tonal template, the variations of the position of the peak will reflect whether this one is directly linked to a syllable (or mora) or not.

4. CONCLUSION

Using this representation, we can extract from the acoustic shape what should be transcribed as lexical value variation and as intonative variation. Thus, it could be used for transcription of tone in field investigations. Furthermore, the intra/inter speaker variability and the adequacy of this model for our corpus lead us to question the association between a tone and a mora, and to assume that the melodic tier is independent of the segmental tier, so that we may not need association lines in phonological representations of tone or accent in this kind of langage.

REFERENCES

[1] DI CRISTO, A. (1978), "*De la microprosodie à l'intonosyntaxe* ", Th. d'Etat, Université de Provence, Aix.
[2] HIRST, D. (1983), "Structures and categories in Prosodic Representations", *Prosody, Model and Measurements*, Cutler & Ladd, Springer, Heidelberg, 93-109.
[3] HIRST, D. (1988), "Tonal Units as Constituents of Prosodic Structure: the Evidence from French and English Intonation", *Autosegmental Studies on Pitch Accent*, Van der Hulst & Smith, Foris, Dordrecht, 151-165.

[4] JOUANNET, F. (1985), "*Prosodologie et Phonologie non linéaire*", GERLA-SELAF, Paris.
[5] JOUANNET, F. (1989), "*Modèles en tonologie (Kirundi et Kinyarwanda)*", Editions du CNRS, Paris.
[6] KIMENYI, A. (1976), "Tone Anticipation in Kinyarwanda", *Studies in Bantu tonology*, SCOPIL 3, Hyman (ed.) 169-181.
[7] ROSSI, M. (1971), "Le Seuil de Glissando ou Seuil de Perception des Variations Tonales pour les Sons de la Parole", *Phonetica*, 23, 1-33.

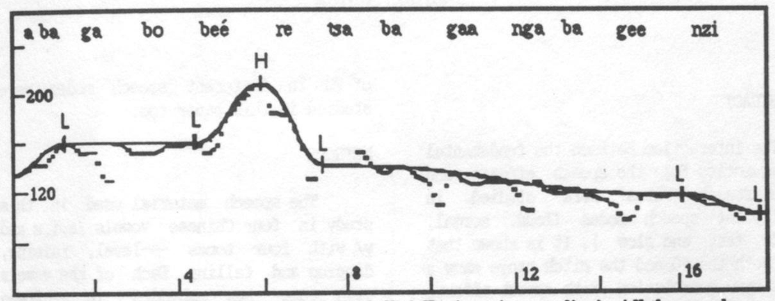

Fig.1: Detected Fo (dotted line) and modelled Fo (continuous line). All the words, except the verb, bear the [LL] tonal template and are produced within one I.U.

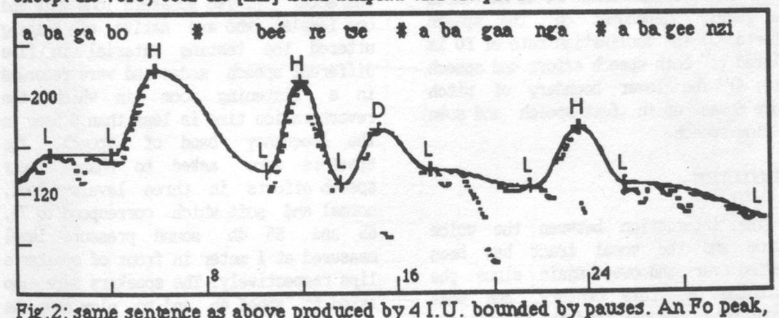

Fig.2: same sentence as above produced by 4 I.U. bounded by pauses. An Fo peak, downstepped if preceded by another one, appears at the end of each I.U.

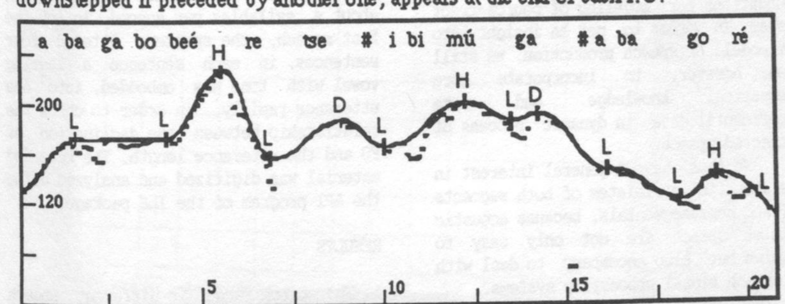

Fig.3: sentence produced by 3 I.U.: a L rises in relation to the distance between the two surrounding peaks while the second one is downstepped. Note the anticipation of the H on /abagoré/, triggered by the [LL] template of the I.U.