



Acoustic cues signaling prosodic units in Moore: a comparison of journalist and non-journalist realizations.

Laetitia Compaoré

Univ Paris Diderot, Sorbonne Paris Cité, LFL, UMR 7110 CNRS, F-75205 Paris, France

laetitia.compaore@etu.univ-paris-diderot.fr

Abstract

This study describes how prosody is used to organize oral speech in Moore (a tone language spoken in Burkina Faso). The analysis pertains to the phonetic realization of intonation in Moore spontaneous speech of two social groups: journalist and non-journalist native speakers of Moore. The main issues are: to identify the acoustic cues which permit speech division into chunks, and to find out if these acoustic indices are the same for the two social groups.

Following the approach of Martin [1], I also consider that prosodic structure relies on the existence of prosodic events instantiated by prosodic contours. Therefore, the prosodic description will determine the acoustic indices which signal prosodic events, often located at units' boundaries. The study focuses on two prosodic units, terminal and non-terminal utterances. The corpus of the study is made up of recordings of three journalists and three native's speakers of Moore.

Results of the analysis suggest that prosodic boundaries in Moore are determined by the combination of three acoustic parameters: final syllable lengthening, F0 contour, and the duration of pause following the boundary. Depending on the communication style, the use of one of the parameters is more frequent.

1. Introduction

In many languages, journalist speech appears as a standard, the language variety used for public communication. Sometimes the standard variety is so widespread that it may be associated with the common way of speaking.

In this study, I make a comparison between the prosody of standard Moore (journalist speech) and that of the layman. How do the two varieties of language use prosody to divide the stream of speech into phrases?

I assume that there are prosodic boundaries in Moore and that the speakers of Moore use them conveniently.

Hypothesis:

- A particular intonation contour is used as a marker of each type of prosodic boundary (terminal and non-terminal) in the two variety of speech.
- Duration of the last two syllables of the prosodic unit depends on the type of prosodic boundary in each style of speech.
- Duration of pauses following the phrases depends on the prosodic boundary type in both journalist and non-journalist speech.

2. Method

2.1. Corpus

There are two types of corpora. The first type is made up of recordings taken from a radio program call « Rites et Coutumes ». The program is broadcast on RTB (Radio Télévision du Burkina). The aim of the program is to go and meet villagers in order to discuss with them and discover the origin and customs of the village. It consists very often in an interview of the village chief, conducted by a journalist.

Extracts of three journalists' speech have been analyzed. The three of them are native speakers of Moore. They also speak French very fluently as French is the official language used for education and by the administration in Burkina Faso. English is a third language for the three journalists.

The second corpus consist of recordings of three native speakers of Moore. One of the subjects is bilingual, she speaks Moore and French (French has been acquired early in her childhood). The other two have an elementary level of French, they cannot carry on a conversation.

Table 1. *Socio-linguistic characteristic of non-journalist speakers.*

Subjects	Age	Occupation	Languages spoken
Subject 1	65	Retired assistant	Mooré; French
Subject 2	44	Housewife	Mooré; French
Subject 3	20	Housekeeper	Mooré; French

The first corpus was taken directly from the radio and the second was recorded with an Edirol R09 compact digital recorder, in a common house. Sounds have been converted into wav format and a mono channel at a sample rate of 22 050 Hz.

The two corpora are made up of spontaneous speech if we refer to [2]: « spontaneous speech ... is a speech that is not read from script ». But we can add that journalist speech is prepared because the subject is known in advance.

106 utterances have been analyzed. They consist of 59 utterances taken from journalist speech and 47 utterances from non-journalist speech.

2.2. Transcription

The first step of the analysis was to listen and to make a phonetic transcription of the corpus. Then using WinPitch [3] (a software which allows a reliable representation of

the fundamental frequency (F0) contour and relying on listening the corpus has been segmented into units. The two units analyzed suggested either completion or non-completion. The units are not necessary minimal prosodic units.

3. Results

3.1. Intonation contour of journalist speech

The analysis of journalist speech first focuses on the intonation contour, displayed by F0. The acoustic analysis shows that the journalist uses an F0 rising contour at the ends of non-terminal prosodic units, whereas the terminal units are associated with a falling contour.

Figure 1 displays the F0 contour of a non-terminal unit produced by a journalist. Over the course of the phrase, the fundamental frequency does not show many variations. It represents a flat intonation contour until the end of the units where the subject makes a sharp rising in the fundamental frequency. The rising is all the more marked as a small falling contour precedes it. Thus, the F0 contour analysis of journalist non-complete utterance indicates that the main prosodic event occurs at the end of the phrase. And a sharp rising contour accompanies generally non-terminal units.

Figure 2 is an example of a terminal utterance produced by a journalist. This utterance consists of two prosodic units. The second unit starts with a little reset of the fundamental frequency. Since this study aim to find out indices signaling a completion, the end of the phrase gives the most insight. Therefore, journalist terminal units display a final falling contour.

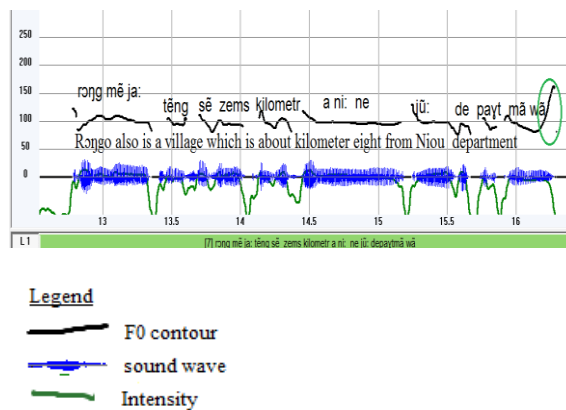


Figure 1: F0 of journalist non-terminal utterance.

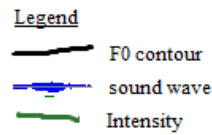
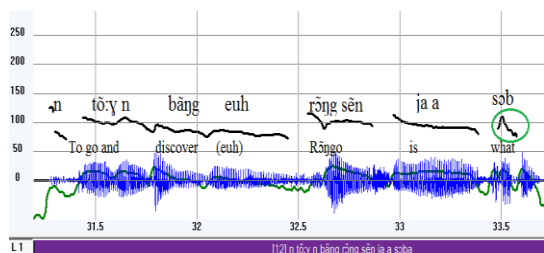


Figure 2: F0 of journalist terminal utterance.

The acoustic analysis of journalist speech has shown F0 contours accompanying terminal a non-terminal phrase. A statistical analysis will allow us to check if these contours are most used in this type of speech. Three simple contours (rising, falling and level) and three complex contours (complex rising, complex falling and complex level) were identified in the corpus. A complex contour is a combination of two melodic movements; the final melodic movement is used to refer to the contour. For example, complex rising end with a rising contour.

The statistical analysis shows that 95% of journalist non-terminal phrases end with rising intonation pattern. Among these rising intonation pattern, 65% of the phrases have a final rising contour, the other 30% end with a complex rising contour.

Beside, 64% of journalist terminal phrases are marked with a falling intonation pattern (falling and complex falling) and 36% of these units terminate with a level contour. So, we can observe that journalists mostly select a falling intonation pattern at the end of complete utterances, even though a level contour can also be used

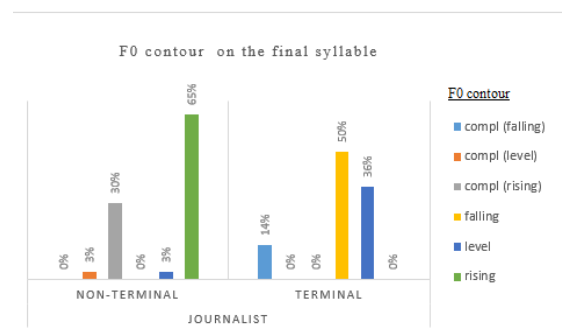


Figure 5: F0 contour on the final syllable.

3.2. Intonation contour of non-journalist speech

The acoustic analysis of Moore common speech also suggests that non-terminal phrases have a final rising intonation pattern and falling contour is associated to terminal phrases.

Figures 3 and 4 are examples of non-terminal and terminal utterances produced by a non-journalist. We can see from figure 3 that the non-terminal utterance produced by common speakers of Moore display more variations in F0. And a complex rising contour (fall rising) accompanies the phrase. Compared to the journalist speech, the rising contour is less marked.

Figure 4 is a terminal utterance of a non-journalist speech. There are also variations of the fundamental frequency of

the entire utterance, and we notice that a falling intonation pattern accompanies terminal phrases.

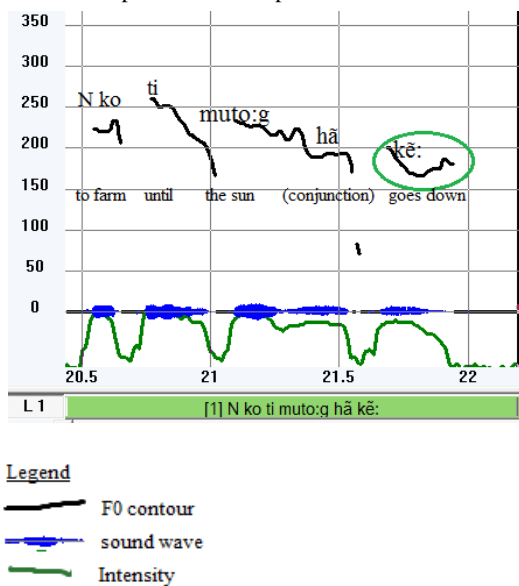


Figure 3: F0 of a non-terminal utterance produced by non-journalist.

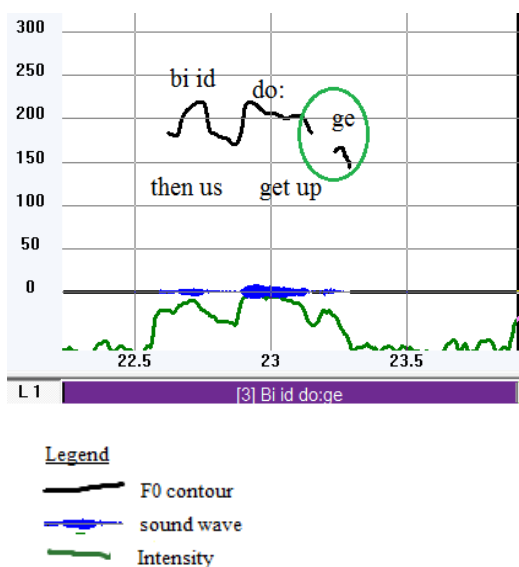


Figure 4: F0 of a terminal utterance produced by non-journalist

I conducted a statistical analysis again to check if the intonation patterns that seem to be associated with non-journalist terminal and non-terminal utterances, are the most used. The analysis indicates that 58% of non-terminal utterances terminate with a rising intonation pattern and 13% of these phrases have level intonation contour. The remaining 29% of non-terminal phrases occur with a falling contour.

Looking at the intonation contour of non-journalist terminal phrases, we can see that 56% of these phrases have a level contour and 33% a falling contour. 11% of

terminal phrases produced by non-journalist terminate with a rising contour. From this analysis, we can conclude that non-journalist does not actually associate a single intonation pattern to a type of phrase. They use different intonation contour for the same type of phrase.

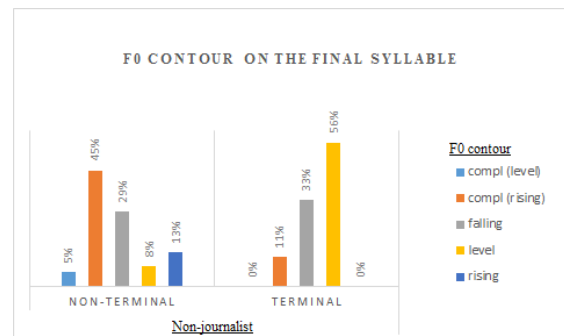


Figure 6: F0 contour on the final syllable.

3.3. Duration of the last two syllables

The second acoustic cue considered in this study is the duration of the last two syllables of the phrase. I posit, after listening to the corpus, that non-terminal phrase boundary is marked with a lengthening of the last syllable; and last syllable of terminal phrase boundary is reduced. Do journalist and non-journalist use durations to distinguish the two type of phrase boundaries? Since duration is a relative value, I make a comparison of the duration of the last two syllables to determine if the final syllable is lengthened or reduced.

In journalist and non-journalist speech, the comparison shows that final syllables of non-complete utterances are on the average longer than the penultimate syllable. Journalists produces 191 ms average duration for penultimate syllable whereas the last syllable is 305 ms average. Beside non-journalists penultimate syllable is 197 ms average and 335 ms duration average of the final syllable.

Looking at the duration of terminal syllable we can see that both groups, journalist and non-journalist produces the final (journalist: 181 ms; non-journalist: 257 ms) a little shorter than the penultimate ones (journalist: 199 ms; non-journalist: 260 ms).

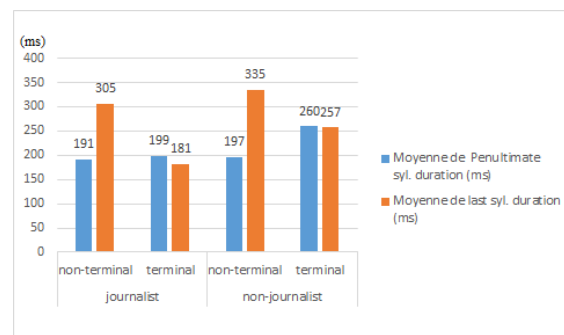


Figure 7: Comparison of final and penultimate syllable duration

A statistical paired t-test has been carried to find out if the differences between the compared means are relevant and not due to chance. The t-tests of the four comparisons indicate that all the differences observed are significant with a $p < 0.05$.

Table 2. *T-test results*

	Journalist	Non-journalist
Terminal	Penultimate syllable duration, last syllable duration Ecart moyen DDL t p -56.227 21 -2.426 .0244	Penultimate syllable duration, last syllable duration Ecart moyen DDL t p -134.556 8 -5.110 .0009
	Non-terminal	Penultimate syllable duration, last syllable duration Ecart moyen DDL t p 70.514 36 4.390 <.0001

3.4. Duration of pauses

The last parameter considered is the duration of pauses. The hypothesis in this analysis is that the duration of pauses depends on the type of the prosodic boundary preceding it. Thus, the main point is to find out how two groups of speakers use this parameter. The results show that two groups use the duration of pauses in the same way. They produce longer pauses after terminal units, almost double the pause duration preceding non-complete units.

Table 3. *Average duration (ms) of pauses after terminal and non-terminal units*

	terminal	Non-terminal
Journalist	715	445
Non-journalist	770	488

4. Discussion

The analysis shows that both groups, journalist and non-journalist use the three acoustic parameters (F0 contour, syllable duration and pauses duration) with little difference.

The duration parameter of syllables and pauses was used in the same way by each group. Final syllables of non-complete units are lengthened while terminal units' last syllables are slightly reduced.

Like pointed out in [4], this study shows that the final lengthening occurs at prosodic boundaries and it is a cue for the prosodic organization of the speech. The lengthening may occur at discourse final utterance boundary or utterance boundary followed by another utterance in the same discourse. However, it is interesting to note that in Moore, the lengthening suggests non-completion, and therefore it is not found with discourse final utterance.

In addition, Pauses following terminal units are produced longer than those following non-complete phrases. These results suggest a rhythmical relation between the last two

syllables and the pauses. When the phrase final syllable is lengthened, the duration of the penultimate syllable and the pause are shortened; and when the final syllable is slightly reduced, the pause and the penultimate syllable are longer.

The results indicate that journalists and non-journalists do not use F0 contour in the same way. Journalist speech do not display many variations of F0 until the end of the phrase. And there is a clear difference between F0 contour associated to terminal units (rising) and non-terminal units (falling). It is also interesting to note that pitch contour used by journalists follows the universal of intonation stated in [5] according to which a rising pitch is associated with non-completion and a falling pitch with completion.

On the other hand, non-journalist speech displays many variations in F0 over the course of the utterance. But there is no clear difference between the types of F0 contour associated to phrases. Many intonation patterns occur with the same type of unit.

The explanation for this contour is that, journalists endeavor to make themselves understood so they use all the three parameters including F0 contour to distinguish their phrases, whereas non-journalists mostly rely on the rhythmic parameter to distinguish their phrases

5. Acknowledgements

I wish to thank Professors Ph. Martin and G. Boulakia for their valuable advice.

6. References:

- [1] Martin, P. (2009). *Intonation du français*. Paris: Armand Colin.
- [2] Beckman, M. E. (1997). A Typology of Spontaneous Speech; in: Sagisaka, Y.; Campbell, N; Higuchi, N. (Eds), *Computing Prosody. Computational Models for Processing Spontaneous Speech*, Springer, 7-26.
- [3] WinPitch www.winpitch.com
- [4] Hofhuis, E., Gussenhoven, C. & Rietveld, A. (1995). Final lengthening at prosodic boundaries in Dutch. *ICPhS 13*. Stockholm: KTH and Dept Linguistics, University of Stockholm. 154-157
- [5] Connell, B. (2005). Pitch realization of questions vs statements in Mambila. In Akinlabi, A. (ed.) *Proceedings of the 4th World Congress of African Linguistics*. Köln: Rüdiger Köppe Verla. 107-116