



Is there a prosodic difference between left-dislocated and heavy subjects? Evidence from spontaneous speech

Mathieu Avanzi, Cédric Gendrot, Anne Lacheret

► To cite this version:

Mathieu Avanzi, Cédric Gendrot, Anne Lacheret. Is there a prosodic difference between left-dislocated and heavy subjects? Evidence from spontaneous speech. Every language, Every Style, Speech Prosody, 5th International conference, 2010, Chicago, United States. pp.20-30. halshs-00637637

HAL Id: halshs-00637637

<https://shs.hal.science/halshs-00637637>

Submitted on 2 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Is there a prosodic difference between left-dislocated and heavy subjects? Evidence from spontaneous French

Mathieu Avanzi^{1,3}, Cédric Gendrot², Anne Lacheret-Dujour³

¹Chaire de linguistique française, Université de Neuchâtel, Neuchâtel, Switzerland

²LPP, UMR 7018/CNRS Université Paris 3-Sorbonne Nouvelle, Paris, France ;

³MoDyCo, Université Paris Ouest Nanterre, Paris, France

mathieu.avanzi@unine.ch, cgendrot@univ-paris3.fr, anne@lacheret.com

Abstract

This paper investigates the prosodic difference between two types of subject-NPs in spontaneous French: left-dislocated NPs (NPs followed by a coreferent pronoun in the subsequent verbal clause, as in *mon mari_i, il_i est instituteur*, *my husband_i he_i is a teacher*) and heavy NPs (NPs which are not followed by a coreferent pronoun, as in *mon mari est instituteur*, *my husband is a teacher*). In order to verify a potential prosodic difference between these two kinds of subject in spontaneous French, two instrumental analyses were conducted: (i) an automatic prominence detection determined whether a boundary tone ends the NP; (ii) a comparison of several acoustic features allowed for the acoustic estimation of the degree of prominence of the NPs of each class. We show that both types of subject-NPs cannot be statistically distinguished according to the measurements on the subject-NPs' final syllable, even if a tendency towards higher prominence values (detection and degree) for dislocated subjects clearly appears. A great variability within and in-between utterances is observed and is suggested to account for the non-significant differences.

Index Terms: clitic left dislocation, heavy subject, prominence, spontaneous French.

1. Introduction

The aim of this study is to provide a clear diagnosis for prosody of left-dislocation in French. In concrete terms, we will try to answer to the following question: is there a prosodic difference between *mon mari est instituteur* and *mon mari_i, il_i est instituteur*? In the first section, we briefly present the literature on this topic. In the second section, we present our corpus and the computational tools for signal processing and prosodic tagging we used. The third and final section, before discussion and conclusion, is devoted to the presentation of the analyses and their results: the prosodic difference between left-dislocated and heavy subject is approached in two steps. First, we focus on the presence or the absence of a boundary tone at the end of the NP. Second, we detail the acoustic correlates of the break's degree of prominence between the NP and the adjacent clause.

2. Previous work

Two ways to diagnose the difference between left-dislocated and heavy subjects in spoken French have been proposed in the literature.

2.1. The difference lies in the (non-)obligatory presence of a boundary tone at the end of the NP

In the Prosodic Theory framework (Selkirk [1]; Delais-Roussarie *et al.* [2] [3]), the two kinds of NPs do not manifest

the same structural properties. While heavy subjects can be ended by a boundary tone or not (the realization of such a tone depending in this case on the speech rate and on the NP's metrical structure: the longer the NP is and the slower the rate of speech is, more the NP will be susceptible to carry a final prominence), left-dislocated subjects are always followed by a major prosodic boundary, that is to say they obligatorily form an Intonational Phrase (IP) on their own. Compare:

- (1) (mon mari est instituteur)_{IP}
- (2) (mon mari)_{IP} (est instituteur)_{IP}
- (3) (mon mari)_{IP} (il est instituteur)_{IP}
- (4) *(mon mari il est instituteur)_{IP}

According to this framework, the first three examples (1)-(3) are possible, while the last one (4) is ungrammatical: obligatory prosodic boundary realization is due to the left-dislocated NP syntactic status (they are adjuncts, Rizzi [4]) and assume a specific pragmatic function (they play the role of topic, Lambrecht [5]). We find the same idea, formulated in different terms, in other frameworks, such as in Martin [6]'s phono-syntactic model or Mertens [7]'s morphological approach.

2.2. The difference lies in the prominence degree of the boundary tone at the end of the NP

Another analysis has been done by Bally [8], one of the first scholars interested in the prosody of spontaneous French. According to this scholar, the difference between the two kinds of NPs lies in the degree of the boundary tone: the break between the subject and the adjacent verbal clause is weak in the first case, strong in the second one. He tried to characterize this shade by saying that in the dislocated utterance the pitch profile is "segmented" because interrupted by a silent pause and by slope inversion, while in the heavy subject utterance the profile is "continuous".

Rossi [9]'s conclusions confirmed [8]'s first perceptual intuitions. [9] compared the prosody of left-dislocated and heavy subjects in read aloud sentences. His results showed that the difference between heavy and left-dislocated subjects relies in the degree of prominence of the terminal syllable of the NP. More precisely, he concluded that the last syllable of the NP seemed to be more salient when the subject was dislocated than when it was heavy. Indeed, he showed that classic prosodic parameters implied in the realization of French stress were more important at the end of a dislocated subject than at the end of a heavy subject. Regarding the last syllable of the left-dislocated element, his measures showed that (i) the melodic target was higher (it reached the highest level of the speaker range), (ii) the rise was broader due to a greater duration (this syllable lengthening entailing the perception of a silent pause).

2.3. Summary

Scholars agree that heavy and left-dislocated subjects express prosodic differences in the way they are prosodically phrased. The problem is that these conclusions have been obtained from the analysis of read aloud speech, that is to say on punctuated texts which have certainly strongly influenced the prosodic phrasing of the subjects (a comma after the left-dislocated NP suggesting a “segmented” prosodic realization; the absence of a comma after a heavy subject facilitating a “continuous” prosodic realization). As far as we know, specific studies addressing the specific question of the acoustic difference between heavy and left-dislocated subjects in spontaneous French do not yet exist (nevertheless, see De Cat [10] for an attempt). On the basis of a certain number of utterances, such an analysis should also allow for the estimation of the prosodic characteristics which would differentiate left-dislocated and heavy subjects. In the next section, we present the data and the instrumentation we used to conduct such an analysis.

3. Material

3.1. Corpus

If it is impossible to obtain minimal pairs when we work on spontaneous speech, we can nevertheless control a certain number of variables, and elaborate a corpus as homogenous as possible. For this purpose, we investigated two spoken Parisian French corpora (a subpart of the PFC (*Phonologie du Français Contemporain*) corpus, 16 speakers, 7 h. of free and semi-directed speech, see Durand *et al.* [11]; the whole CFPP2000, 22 speakers, 17 h. of semi-directed speech, see Branca *et al.* [12]) and manually extracted only subject-NPs which comprised 2 or 3 syllables. We collected altogether 134 utterances, namely 67 left-dislocated subjects and 67 heavy subjects. Each pair was pronounced by the same speaker. Let us note that not all speakers do alternate between dislocated and heavy subjects; moreover, there are speakers who frequently alternate, while others do not. That’s why our corpus contains only 18 speakers (13 females and 5 males, from 24 to 85 years old) out of the 38 speakers the whole corpus contains. These 18 speakers perform at least 1, at most 11 items. To conclude on this point, we would like to highlight that only utterances where the pronominal recovery has been made with a verb pre-posed personal subject pronoun, such as *je, tu, il/elle/on, nous, vous, ils/elles*, have been considered. This means the utterances where the recovery was done with a demonstrative pronoun, such as those from the */c-/* paradigm (like *ce, cela, ça, ceux-ci, etc.*), had not been taken into account because they give the NP a particular semantic status (*i.e.* the NP gains a generic interpretation, and the sentence would in turn be like a pseudo-cleft-sentence, see Carlier [13], Pekarek-Doehler & Müller [14]) and they do not resort to the same syntactic analysis (Avanzi [15] [16] for the details of the argumentation).

3.2. Prosodic preprocessing

All the utterances of the corpus (2 x 67 utterances) were aligned semi-automatically into phones, syllables and orthographic words with the EasyAlign (Goldman [17]) Praat script (Boersma & Weenick [18]). Next, we conducted prosodic prominence detection with the ANALOR software. Because of lack of space, we cannot present the entire procedure here (see Avanzi *et al.* [19] for a first experiment, [20] for the last state of description of the tool). In a few words, the algorithm calculates the relative height and duration of each syllable in a given utterance by comparing the value of

the analyzed syllable with the average of the six adjacent syllables (*i.e.* the three preceding and the three following ones); the pitch rise slope is then processed and the presence of a subsequent silent pause is considered. Regarding duration, a simple relativization was carried out. Ideally, syllable duration models based on syllable intrinsic properties and normalizing local speech rate for spontaneous speech should be used (see for a first experiment on read aloud sentences Obin *et al.* [21] [22]). Given that such a model still has to be reliably developed and tested, we only took the number of phonemes of the syllable into account. By doing so, we could avoid the bias of syllable weight, a syllable compounded of five phonemes being by essence longer than a mono-phonemic one. Pitch measures (average height and rise) are in semi-tones.

According to thresholds which were automatically estimated on the basis of the annotation of a 70-minute corpus by two experts (see Simon *et al.* [23] and Avanzi *et al.* [24] for an exhaustive presentation of this material); the ANALOR software program marks the prominent syllables and estimates the perceptual degree of each one. These calculations are described in the following section (see Figure 3 and Figure 4 §8 for illustrations).

4. Analysis and results

This section is focused on the presentation of the prosodic analysis. Considering that in French, prosodic boundaries coincide with prominent terminal phrases’ syllables (Rossi [25], Vaissière [26] and Lacheret-Dujour & Beaugendre [27]), we used the ANALOR software to conduct a prosodic analysis in two steps. First, we searched to see whether dislocated NPs systematically ended on a prosodic boundary, as the Prosodic Theory predicts. Next, we compared the degree of prominence of each syllable of our two classes to see if we found a significant difference between them.

4.1. On the presence of a boundary

In the first step of the study, we used the ANALOR tool to check whether the dislocated NPs of our corpus were always followed by a prosodic break, and compared the results with those for the heavy subjects. We considered that a syllable was prominent or bearing a boundary tone if one of the following four parameters was activated: the relative pitch height reaches at least 2.5 st., relative duration is 1.5 longer, vocalic rise is equal or greater than 2.5 st. and/or a silent pause, whatever its duration, follows the syllable. As one can see in the diagram in Figure 1, contrary to the predictions developed above in 1. a boundary tone is not systematically detected at the end of the left-dislocated NP:

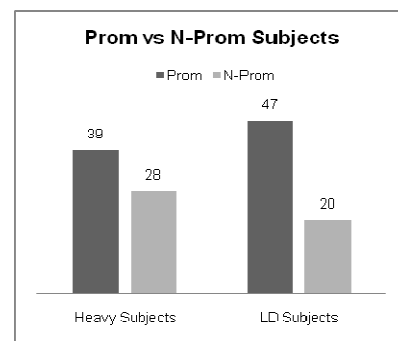


Figure 1: Distribution between the number of Heavy Subjects (on the left) and the number of Left-Dislocated Subjects (on the right) according to the

presence or the absence of a prominence on the terminal syllable of the NP.

A binomial logistic regression was used in order to predict the probability of prominence detection (yes/no) according to subject's type and speaker. This makes apparent that the impact of the subject's type on the prominence detection is not significant ($\text{Chi}^2 = 2.46$, $p=0.11$), and neither is the impact of speaker ($\text{Chi}^2 = 12.6$, $p=0.76$).

4.2. On the boundary degree

A second analysis has been conducted to see if the difference between heavy and left-dislocated subjects had to be searched in the prominence degree of the boundary tone. Estimating the degree of salience of accentual prominence in French is not so easy: prosodic parameters involved in French accentual prominence are numerous, and there are many compensatory phenomena between them (that is to say, for example, that in some contexts, a pitch rise can be equivalent perceptively to a vocalic lengthening, see t'Hart *et al.* [28], *inter alia*). Nowadays, devoted tools that can deal with it are still missing. In this study, we propose a method to estimate the degree of prominence of a syllable. Our study is based upon the following hypothesis: the more the acoustic parameters are mobilized (number of parameters and/or prominence thresholds), the more the prominence will be perceived. The thresholds mentioned above (§4.1.) have been used to assign a score to each syllable. 1 was given to a syllable which had an approaching threshold (25% less than the chosen threshold), 2 if the value reached but not exceeding 25% of the original threshold, 3 if the threshold exceeded more than 25% of the reference measure. We attributed 0 if the threshold did not reach at least 25% less than the training value; 1 if the syllable was followed by a silent pause. This allowed for an *ad hoc* non linear rescaling of prominence between 0 and 10 (in most 3 for relative height, 3 for duration, 3 for pitch rise and 1 for silent pause). The calculation had been applied to all utterances of the corpus, which gave the following distribution:

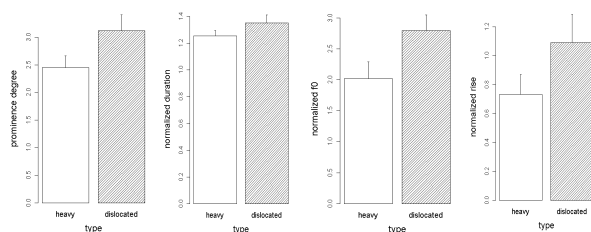


Figure 2: Degree of prominence; normalized duration; normalized f0 and normalized f0 rise (from left to right) for heavy vs. dislocated subjects.

A linear mixed effects model with crossed random effects was run, with the “degree of prominence” as the dependent variable and subject's type as a predictor. Speakers were entered as a random term. The advantage of a mixed effects model here is that it takes into account the randomness induced by the different number of realizations produced by each speaker but also their variations. The model shows no effect of subject's type on the prominence degree ($F(1,132) = 3.75$, $p = 0.055$), however it does show a tendency towards higher degree of prominence for dislocated subjects. Separate measurements of normalized duration, normalized f0 and normalized f0 rise displayed in Figure 2 also follow the same tendency. The standard error bars give a hint about why this tendency doesn't reveal significant differences. A great variability within and between speakers is observed for the

production of both subject-NPs. Furthermore 6 out of 18 speakers have higher values for dislocated subjects than for heavy subjects. This variability is a strong clue to invalidate [8]'s intuitions and [9]'s studies on spoken aloud utterances.

5. Discussion & Conclusion

The aim of this study was to answer the following question: is there a difference regarding the prosodic boundary between heavy and left-dislocated subjects. In order to bring new arguments, we analyzed a spontaneous corpus of 67 pairs of utterances pronounced in semi-directed conditions by Parisian French.

Contrary to the traditional predictions (see §2. above), the first analysis showed that a strong prosodic break is not obligatory at the end of left-dislocated subjects. The second analysis showed that the breaks which end the left-dislocated NPs are slightly more prominent than the breaks which end the Heavy NPs, but that the difference is not significant, for neither concerning the “global” score, nor any of the acoustic parameters investigated (relative height, relative duration and intra-vocalic pitch rise). It also revealed that two strategies were used by the Parisian speakers of our corpus. One consists of making a greater prominence on the left-dislocated NP (12/18 speakers); the other consists of making a greater prominence on the heavy-subject NP (6/18 speakers). While the first can be explained by taking syntactic information into account (extra-sentential elements are always ended by a boundary tone), the second can be explained by an economy principle (Lacheret-Dujour [29] and [30]), which is relatively close to Gussenhoven [31]'s effort code. Applied to the utterances we studied here, it stipulates that a right edge of a discourse constituent need not be marked at the same time by syntax and by prosody if one of the two parameters already operates. That is to say that a subject NP need not be prosodically salient if the syntax already marks it as dislocated, and on the contrary, a subject NP already prosodically salient is not obligatory marked by the syntax.

These findings deserve further consideration. (i) We will have to integrate speech rate and the syntactic position of the NP in the sentence as predictor variables, the sentences pronounced with a faster speech rate will tend to contain fewer prominences, non sentence initial NPs (embedded clauses, NPs following a connector) do not have the same metrical structure as initial ones. (ii) We will also have to see if, for a certain number of cases, the absence of a detected prosodic break is not a bias from the software. Indeed, the actual algorithm does not allow us to take other prosodic characteristics, such as the low tones, into account. As [6] and Le Gac & Yoo [32] recently argued, such realizations exist: they are motivated by slope inversion. (iii) From the informational structure point of view, we will have to cross the results we get in this study with a categorization of topics, and see how syntactic dislocation and prosodic detachment interact. We will focus our attention on these three points in future work.

6. Acknowledgements

Two institutional frameworks financed this research: The Swiss National Science Foundation (SNF Project 100012-113726/1, “La structure interne des périodes”, Neuchâtel University), and the Agence Nationale de la Recherche (ANR-07-CORP-030-01, “Rhapsodie – Corpus prosodique de référence du français parlé”). Thanks also to B. Victorri (Lattice/ENS, Paris) for friendly help in the different stages of the ANALOR software conception.

7. References

- [1] Selkirk, E. *Phonology and syntax. The relations between sound and structure*, MIT Press, 1984.
- [2] Delais-Roussarie, E. Doetjes, J. & P. Sleeman, "Dislocation", in Corblin, F. & H. de Swart (eds.), *Handbook of French Semantic*, Stanford, 501-528, 2004.
- [3] Delais-Roussarie E. & Post, B. "Unités prosodiques et grammairale de l'intonation: vers une nouvelle approche", *Actes des Journées d'étude sur la Parole JEP-TALN 08*, 2008.
- [4] Rizzi, L. "The fine structure of left periphery", in Haegman, L. (ed.), *Elements of Grammar. Handbook in generative syntax*. Dordrecht, Kluwer, 281-337, 1997.
- [5] Lambrecht, K. *Topic, antitopic, and verb agreement in non-standard French*, Amsterdam, John Benjamins, 1981.
- [6] Martin, Ph. *L'intonation du français*, Paris, Armand Colin, 2009.
- [7] Mertens, P. "A Predictive Approach to the Analysis of Intonation in Discourse in French, in Kawaguchi, Y. Fonagy, I. Moriguchi, T. (eds), *Prosody and Syntax*, Amsterdam, John Benjamins, 64-101, 2006.
- [8] Bally, C. *Linguistique générale et Linguistique française*, Berne, Francke, 1944.
- [9] Rossi, M. *L'intonation, le système du français – description et modélisation*. Paris/Gap, Ophrys, 1999.
- [10] De Cat, C. *French Dislocation. Interpretation, Syntax, Acquisition*, Oxford, University Press, 2007.
- [11] Durand, J. Laks, B. & Lyche, Ch. "La phonologie du français contemporain: usages, variétés et structure", in Pusch, C. & W. Raible (eds), *Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen, Gunter Narr Verlag, 93-106, 2002.
- [12] Branca-Rosoff, S. et al. *Discours sur la ville. Corpus de Français Parlé Parisien des années 2000 (CFPP2000)*. <http://ed268.univ-paris3.fr/CFPP2000/>, 2009.
- [13] Carlier, A. "Les gosses ça se lève tôt le matin": l'interprétation générique du syntagme nominal disloqué au moyen de *ce* ou *ça*", *Journal of French Language Studies*, 6, 133-162, 1996.
- [14] Pekarek-Doehler, S. & G. Müller, "Le problème c'est de les distinguer. Disloquée à gauche et pseudo-clivée dans la conversation", in Apothéloz, D. et al. (eds), *Les linguistiques du détachement*, Bern, Peter Lang, 413-426, 2009.
- [15] Avanzi, M. "La dislocation à gauche avec reprise anaphorique en français parlé. Etude prosodique", *IDP09 Proceedings*, 2009.
- [16] Avanzi, M. "La dislocation à gauche en français. Etude instrumentale". *Le français moderne*, 2011 (2).
- [17] Goldman, J.-P. "EasyAlign: a semi-automatic phonetic alignment tool under Praat", <http://latcui.unige.ch/phonetique>, 2008.
- [18] Boersma, P. & Weenink, D. *Praat: doing phonetics by computer (Version 5.1)*. www.praat.org, 2009.
- [19] Avanzi, M. Lacheret-Dujour, A. & Victorri, B. "ANALOR. A Tool for Semi-Automatic Annotation of French Prosodic Structure", *Proceedings of Speech Prosody'08*, 119-122, 2008.
- [20] Avanzi, M. Lacheret-Dujour, A. & Victorri, B. "A corpus-based learning method for prominence detection in spontaneous speech", *Prosodic Prominence: Perceptual and Automatic Identification (Speech Prosody 2010 workshop)*, Chicago, USA, 2010.
- [21] Obin, N., Rodet, X & A. Lacheret-Dujour, "Un modèle de durée des syllabes fondé sur les propriétés syllabiques intrinsèques et les variations locales de débit", *Actes des Journées d'étude sur la Parole JEP-TALN 08*, 2008.
- [22] Obin, N, Rodet, X., & A. Lacheret-Dujour, "A Multi-Level Context-Dependent Prosodic Model Applied to Durational Modeling", *Interspeech*, Brighton, UK, 2009.
- [23] Simon, A.C. Avanzi, M. & Goldman, J.-P. "La détection des prééminences syllabiques. Un aller-retour entre l'annotation manuelle et le traitement automatique", *Actes du 1^{er} Congrès Mondial de Linguistique Française (CMLF 08)*, 1673-1686, 2008.
- [24] Avanzi, M. Goldman, J.-P. Simon, A.C. & A. Auchlin "C-PROM. An Annotated Corpus for French Prominence Studies", *Prosodic Prominence: Perceptual and Automatic Identification (Speech Prosody 2010 workshop)*, Chicago, USA, 2010.
- [25] Rossi, M. "Le français, langue sans accent ?", *Studia Phonetica*, 13-51, 1979.
- [26] Vaissière, J. "Rythm, accentuation and final lengthening in French", in Sunberg et al. (eds), *Music, Language Speech and Brain*, Macmillan Press, 108-120, 1990.
- [27] Lacheret-Dujour, A. & Beaugendre, F. *La prosodie du français*, Paris, CNRS, 1999.
- [28] t'Hart, J. Collier, R. & A. Cohen (1991). *A Perceptual Study of Intonation. An Experimental-Phonetic Approach to Speech Melody*, Cambridge, University Press.
- [29] Lacheret-Dujour A. "The Intonational Marking of Topical Saliency in Spontaneous Speech – Evidence from Spoken French", *Speech Prosody 2002 Proceedings*, 447-450, 2002.
- [30] Lacheret-Dujour, A. & J. François, "De la notion de détachement topical à celle de constituant thématique extrapositionnel". *Cahiers de praxématique*, 40, 167-198, 2004.
- [31] Gussenhoven, C. "Intonation and interpretation: phonetics and phonology", *Speech Prosody 2002 Proceedings*, 47-57, 2002.
- [32] Le Gac, D. & Yoo, H. "De l'invariante invariabilité des topiques en français", *IDP09 Proceedings*, 2009.

8. Annexes

The figures below are ANALOR screen shots. In the abscissa, temporal values are given in milliseconds; in the ordinate, the values of F0 in a logarithmic scale can be seen. The transcription tiers are, from top to bottom: phone tier, syllable tier (both in SAMPA alphabet); automatic prominence detection tier (prominence characteristics can be posted by clicking on red crosses upside the phone tier) and the orthographic word tier. Figure 3 shows a non-prominent left-dislocated subject, Figure 4 shows a prominent heavy subject. Both utterances are performed by the same speaker (a 60 years-old female living in the 12th district of Paris). The final NP syllables (*les parents*) are framed in red.

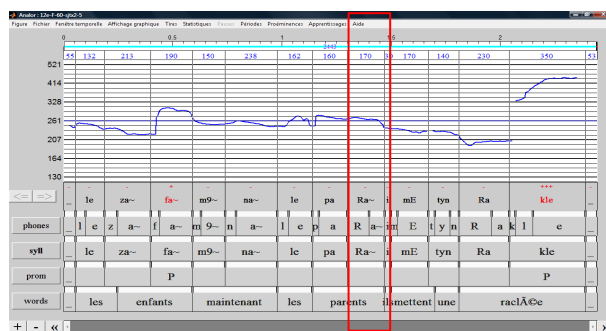


Figure 3: ANALOR screen shot. Transcription of the utterance: (...) *les parents* ils mettent une raclette [cfpp2000, 12-F-60sjtx1-5].

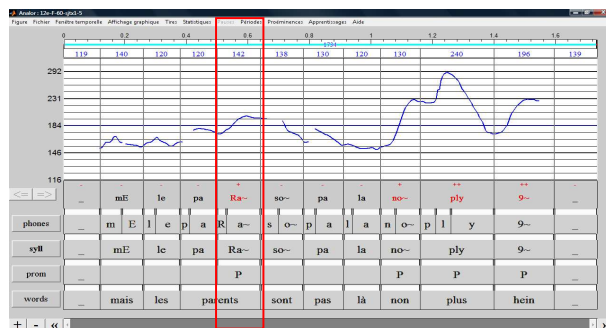


Figure 4: ANALOR screen shot. Transcription of the utterance: (...) *mais les parents* sont pas là non plus hein [cfpp2000, 12-F-60sjtx2-5].