

## Transcribing prosodic breaks in French spontaneous speech: the role of discourse genre and discourse content

Mariapaola D'Imperio\*, Roxane Bertrand\*\*, Cristel Portes\* and Laurent Prévot\*

\*Université d'Aix-Marseille & Laboratoire Parole et Langage, CNRS, France

\*\*Laboratoire Parole et Langage, CNRS, France

[mariapaola.dimperio@lpl-aix.fr](mailto:mariapaola.dimperio@lpl-aix.fr), [roxane.bertrand@lpl-aix.fr](mailto:roxane.bertrand@lpl-aix.fr), [cristel.portes@lpl-aix.fr](mailto:cristel.portes@lpl-aix.fr), [laurent.prevot@lpl-aix.fr](mailto:laurent.prevot@lpl-aix.fr)

Despite a large quantity of studies on French prosody, a consensus transcription system is still needed. This would spur the development of common prosodic annotation strategies for the purpose of data comparison within the research community as well as facilitating automatic annotation. Semi-automatic transcription systems of tonal structure based on either acoustic modeling (INTSINT, Hirst 2005) or on perception-based algorithms (*Prosogram*, Mertens 2004) have been proposed for French. Strictly phonological, manual, approaches to prosodic transcription in French (Hirst & Di Cristo 1998; Jun & Fougeron 2000; Post 2000) are also available. These approaches have however been largely criticized by the French speech community on the basis of a putative deficiency in capturing, among the other things, inter-speaker variability and discourse complexity in spontaneous speech corpora. The point of view supported here is that discourse complexity (disfluencies, interruptions, etc.) does not preclude the use of phonological units for prosodic labeling (cf., Nesterenko *et al.* 2010, Portes & Bertrand 2011). Moreover, phonological approaches have been judged to be largely not applicable for the purposes of automatic prosodic labeling. Hence, more recently, a theory-independent model has been proposed (Lacheret *et al.* 2010). Note however that lack of explicitly defined criteria for category annotation (either prosodic constituents or tonal events) can be an obstacle to the robustness and agreement among transcribers.

A major need is hence to develop a phonologically-based transcription system that would be flexible enough so as to be employed across discourse genres and be amenable to corpus-based learning and automatic labeling. The first step in such an enterprise is to set up an inventory of phonologically defined prosodic units, which can be used as a basis for defining a limited number of break levels. Note that previous attempts to manually or automatically annotated break levels in French only include 2 levels, the Intonation Phrase (IP) and the Accentual Phrase (AP). Our goal here is hence to present a newly devised set of guidelines for transcribing prosodic units in different French speech styles including an intermediate phrase level (*ip*, D'Imperio & Michelas 2010). We also set out to test the potential for non-expert manual annotation based on our criteria. Specifically, we trained 4 non-expert transcribers and then asked them to transcribe 4 levels of prosodic break (0 = no break; 1 = AP break; 2 = *ip* break; 3 = IP break) in 4 sets of conversational data. Prosodic levels were defined and identified in terms of a ToBI-style annotation (both on a perceptual basis and employing the acoustic representation).

In order to test whether discourse genre had an impact in the annotation performance, we drew 2 datasets from a corpus of spontaneous dialogues (CID, Bertrand 2008) and from a political debate speech corpus (Portes 2004), for a total of 4 speech segments (4 speakers) of 2 minutes each. Note that the datasets were extracted from quasi-monological sections of the corpora, in order to avoid interactional effects (feedback, overlap, etc.). Moreover, the influence of discourse content (lexical, semantic, syntactic and pragmatic content) on the segmentation choices was also tested by delexicalizing the 4 speech samples and by proceeding to their transcription. We specifically ensured that each transcriber only annotated 2 lexicalized and 2 delexicalized samples (so that no transcriber annotated both the delexicalized and non delexicalized version of the same dataset). The details of our annotation guidelines, as well as inter-transcriber agreement scores and kappa statistics performed on the non-expert annotation data, will be discussed in the light of competing proposals to prosodic transcription. Finally, we evaluated the performance of existing automatic tools (Blache *et al.* 2009) in reference to a gold standard obtained through manual annotations.

## REFERENCES

- Bertrand R. 2008. Le CID : <[http://crdo.fr/voir\\_depot.php?langue=fr&id=000027](http://crdo.fr/voir_depot.php?langue=fr&id=000027)>
- Blache P., Bertrand R. & Ferré G. 2009. "Creating and exploiting multimodal annotated corpora: the ToMA project". In Kipp M. (eds.) *Multimodal Corpora*. Berlin: Springer-Verlag, vol. LNAI 5509, 16, 38–53.
- D'Imperio, M. & Michelas, A. (2010). "Embedded register levels and prosodic phrasing in French". *Proceedings of the Speech Prosody 2010*, Chicago, USA.
- Hirst, D.J., 2005. "Form and function in the representation of speech prosody". In K.Hirose, D.J.Hirst & Y.Sagisaka (eds) *Quantitative prosody modeling for natural speech description and generation* (=Speech Communication 46 (3-4)), 334-347.
- Hirst & Di Cristo, 1998 *Intonation Systems: A Survey of Twenty Languages*, CUP Press.
- Jun, S.-A., & Fougeron, C. (2000). A phonological model of French intonation. In A. Botinis (Ed.), *Intonation: Analysis, modelling and technology*. Boston: Kluwer.
- Lacheret, A., Obin, N. & Avanzi, M. (2010). "Design and Evaluation of Shared Prosodic Annotation for Spontaneous French Speech: From Expert Knowledge to Non-Expert Annotation", *Linguistic Annotation Workshop*, p. 265-273, Uppsala, Sweden, 2010.
- Mertens, P. (2004). "The Prosogram : Semi-Automatic Transcription of Prosody based on a Tonal Perception Model". In B. Bel & I. Marlien (eds.) *Proceedings of Speech Prosody 2004*, Nara (Japan), 23-26 March.
- Nesterenko I., Rauzy S. & Bertrand R. (2010). "Prosody in a corpus of French spontaneous speech: perception, annotation and prosody ~ syntax interaction". *Proceedings of Speech Prosody 2010*, May 11-14 : Chicago, United States of America.
- Portes C. (2004), *Prosodie et économie du discours : Spécificité phonétique, écologie discursive et portée pragmatique de l'intonation d'implication*, Doctoral Dissertation, Aix-Marseille 1, Aix en Provence.
- Portes, C. & Bertrand, R. (2011). "Permanence et variation des unités prosodiques dans le discours et l'interaction". *Journal of French Language Studies*. 2011, 97-110.
- Post, B. (2000). *Tonal and phrasal structures in French intonation* (published PhD dissertation). The Hague: Holland Academic Graphics.