Transcription of Prosody in continuous speech:
Prosodic encoding the Spoken French corpus Rhapsodie

Mathieu Avanzi[1,2,] Anne Lacheret[2,3,] Nicolas Obin[4]
[1]Université de Neuchâtel, Switzerland
[2]MoDyCo Lab., University of Paris Ouest – La Défense, France
[3]IR Corpus Consortium corpus oraux et multimodaux (Fédération de Recherche Typologie et universaux linguistiques), France
[4]IRCAM-CNRS, UMR-9912-STMS, France

In the area of speech processing, there is a demonstrated need for robust prosodic notation systems of continuous speech – regardless to the speaking style (+/- spontaneous vs. formal). The fundamental issue for the transcription of prosody is to define stable events (boundary tones, pitch accents, disfluent segments, etc.) that can be identified both by experts (regardless their theoretical background), and novice listeners (linguists of other domains, students). In particular, the transcription methodology should provide an annotation scheme that can be easily used for prosodic labelling and statistical determination of relevant perceptual-acoustic correlates in continuous speech (Buhman *et al.* 2002; Cresti & Moneglia 2005; Cheng *et al.* 2005; Obin *et al.* 2008; Cole *et al.* 2011, Lacheret-Dujour *et al.* 2011; Smith 2011).

The aim of this communication is to present the different steps of the prosodic transcription study that has been conducted during 3 years within the *Rhapsodie* project in order to provide a reference transcription system, based on prominence labelling as well as disfluencies tagging, for the segmentation of French discourses into prosodic units of different levels. The methodology – including linguistic assumptions, speech database, and experiments results – will be presented as follows:

1. The general context and the different modules involved in the transcription of the *Rhapsodie* speech database, http://rhapsodie.risc.cnrs.fr/fr/, a 33.000 word treebank of spoken French created with the aim of modeling the interface between prosody, syntax and discourse, will be presented, so as the linguistic motivations of the chosen methodology for the prosodic processing (bottom-up approach driven by the data, regardless to a particular theory and functional cues).
2. The speech database, covering various discourse genres and speaking styles (about 3 hours of continuous speech, monologues and dialogues, private vs. public, face-to-face vs. broadcasting, mor or less interactive, descriptive vs. argumentative vs. procedural samples).
3. The results of the experiments conducted to provide guidelines for the transcription of French prosody (methodology and reliability measures) and to design a reference speech database of French prosody.

As for the third point, first: two pilot experiments conducted with a consortium of 15 French experts in order to define the best, *i.e.* the more sharable, transcription unit (boundary vs. prominences) and to process contexts of disfluencies will be resumed. Second: the perception-driven prosodic labelling used in the *Rhapsodie* project, based on these two pilot experiments, and the annotation schemata retained for the transcription will be presented.

References

ANR Rhapsodie 07 Corp-030-01, Corpus prosodique de référence du français parlé, http://rhapsodie.risc.cnrs.fr/fr/index.html.

Buhmann, J., Caspers, J., van Heuven, V., Hoekstra, H., Martens, J.-P., Swerts, M. (2002). "Annotation of Prominent Words, Prosodic Boundaries and Segmental Lengthening by Non Expert Transcribers in the Spoken Dutch Corpus". *Proceedings of LREC 2002*, 779-785.

Cole, J., Mo, Y., Hasegawa-Johnson, M. (2010). "Signal-based and expectation-based factors in the perception of prosodic prominence". *Laboratory Phonology*, 1, 425-452.

Cheng, W., Greaves, C. and Warren, M. (2005). "The creation of a prosodically transcribed intercultural corpus: The Hong Kong Corpus of Spoken English (prosodic)", *International Computer Archive of Modern English (ICAME) Journal*, Vol. 29, 5-26.

Cresti, E., Moneglia, M. (2005). *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam/New York: John Benjamins.

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", *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, 4th Linguistic Annotation Workshop, Uppsala*, 265-273.

Lacheret, A. Kahane, S., Pietrandrea, P. (2013 in preparation). *Rhapsodie: a Prosodic and Syntactic Treebank for Spoken French,* to be submitted (John Benjamins).

Obin, N., Rodet, X., Lacheret-Dujour, A. (2008). "French Prominence: a Probabilistic Framework", *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 3993-3996, Las Vegas, USA.

Smith, C. (2011). "Naïve Listeners' Perceptions of French Prosody Compared to the Predictions of Theoretical Models". *Proceedings of the 3rd IDP Conference*, 349-335.