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/](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


