Some perceptual cues to French prosody
Guillaume Rolland, Hélène Lœvenbruck, and Marie-Agnès Cathiard
Institut de la Communication Parlée – CNRS UMR 5009 – INPG – Université Stendhal, France
rolland@icp.inpg.fr, loeven@icp.inpg.fr, cathiard@icp.inpg.fr

The aim of this study is to examine the perceptual correlates of the lowest intonational unit in Jun & Fougeron’s (2000, in press) model of French intonation, the Accentual Phrase (AP). Two questions are addressed: (1) is the AP a well-identified unit for French listeners? (2) what makes an AP be perceived as such? The AP corresponds to Vaissière’s (1997) ‘prosodic word’, Mertens’s (1993) ‘Intonation Group’, Rossi’s (1985) ‘Intonème Mineur’, Pasdeloup’s (1992) ‘Groupe Accentuel’ and roughly to Di Cristo & Hirst’s (1993) ‘Rhythmic Unit’. It is right-demarcated by the primary stress. Its default tonal pattern is /LHiLH*/, with the initial rise LHi categorized as a phrase accent whose tones are not associated with any syllable, and the H* portion of the LH* pitch accent considered as associated with the AP-final full syllable. Its actual realization depends on the number of syllables, on phonological constraints and rhythmic principles. A corpus of 22 sentences composed of 3 APs (default pattern /LHiLH* LHiLH* LHiL%/) has been recorded for four native French speakers. The prosodic transcription revealed that, when the number of syllables was sufficient, APs were mainly realized [LHiLH*]. A subsequent acoustic analysis showed that the final F0 peak (H*) was generally carried by the AP-final syllable only (rise included) and was usually higher than that of the initial accent (Hi), the initial rise being possibly carried by 2 or 3 syllables. Similarly, the normalized duration of the H* syllable was generally longer than that of the Hi bearing syllable. We then evaluated the ability of French speakers to slice utterances into APs. Ten listeners were presented with the 88 original sentences (4 speakers) as well as with 88 delexicalized sentences in which only original F0, intensity and duration were kept, and syllables were replaced by [ma] using PSOLA. They were asked to choose among 5 phrasings (marked with parentheses) which corresponded best to the utterance heard. Statistical tests show that the listeners were successful in slicing the utterances into APs. Furthermore, they were as good for the delexicalized corpus as for the original one, which means they did rely on prosodic cues to select a phrasing. Among the sentences with a performance score higher than 70%, 78% had one of their first 2 APs realized [LHiLH*]. We suggest that the tonal (low-high-low-higher), durational (AP-final syllable longest) and dynamic (long small rise – short large rise) patterns may be perceptual cues to the AP.