TOWARDS A MODEL OF ITALIAN INTONATION

C Avesani

Scuola Normale Superiore, Italy

This paper presents a model of Italian intonation included in a speech synthesis system developed at Cselt (Torino). The model is based on two theoretical assumptions. The first concerns the structure of the intonational system: following Pierrehumbert (1980) and Liberman & Pierrehumbert (1984), we think that intonation is the result of the interaction of principled choices in four different systems: phrasing, tune, pitch accent placement and pitch range. The second concerns the actual shape of FO contours: we think is determined by categorical choices made from a small set of discrete phonological elements (ladd, 1987).

The system works in two stages: in the first stage, text string is annotated with indications of the relevant intonation events: at this stage a phonological representation of the text is actually performed, and specific tonal elements are properly lined up with the text in order to define phrasing (association of phrase accents and boundary tones), location of prominence points (association of pitch accents to prominent words) and tune (specification of the sequence of pitch accents, phrase accents and boundary tones from a grammar of allowable tonal elements) of the sentence. In the second stage the 'prosodic markers' are assigned to specific segments of the text, converted in target values in a given range, and connected by transition rules. This second stage is the level in which operate a set of specific phonetic rules governing the proper association of certain phonological elements to the text. From several studies on natural sentences and from empirical data on synthetic speech, we discovered that the association point of the tone crucially linked to the stressed syllable vary significantly according to the position of the word in the sentence. Therefore, we stated different linking points for pitch accents occurring in words in sentence-initial, medial, final position.

The system performs a suitable phonological-phonetic representation of simple and complex declarative sentences: it applies also to a special subsets of declarative sentences such as: isolated words; complex NP, typically used as answers in specific contexts; utterances including lists of words.

At sentence level, intonational phrasing is a major aspect of the intonational description given that we have a limited access to morpho-syntactic information, in synthesising complex syntactic sentences, we rely on extra-syntactic informations and on simple pragmatic notions (sentence topic). On these assumptions, we derived two kinds of intonational phrases, whose boundaries are phonetically realized in terms of different sets of intonational and durational properties.

At a discourse level, the model includes systematic variations of pitch range, in order to highlight discourse segmentation and topic structure (Hirschberg & Pierrehumbert, 1986; Avesani & Vayra, 1988).