PROSODIC ANNOTATION OF ORAL ARCHIVES (1913-2018)

PHILIPPE MARTIN
LLF, UFRL, Université Paris Diderot

Oral archives found in newsreels, radio and TV recordings are often characterized by less than optimal recording quality making them unsuitable for prosodic research. In particular, the reduced speech bandwidth due to the then available technology for sound capture and conservation (carbon microphone, vinyl records, optical and magnetic tape sound recording) make intonation analysis difficult and unreliable for most pitch tracking algorithms, such as the ones found in the popular software program Praat. Specific problems are linked to reduced signal intensity, lack of speech fundamental frequency in the spectrum, presence of echo and overlapping voices or sound sources, etc.

To address these limitations and pursue prosodic research on valuable speech archives, a set of dedicated functions have been implemented in the software program WinPitch. These functions are based on the visual validation of the pitch curve against the simultaneous display of a frequency aligned narrow-band spectrogram first (or second) harmonic. Using one of the seven tracking algorithms available in WinPitch (autocorrelation, AMDF, spectral comb, spectral brush, Cepstrum…), the user is not only able to visually check the validity of a given pitch track section against the corresponding spectrogram first harmonic segment, but also to select graphically the F0 curve section of interest and enter its characteristics by graphic commands.

These user-friendly commands using only mouse controls allow to position graphic piecewise linear segments approximating any shape of the fundamental frequency curve. Furthermore, these graphic annotations can be user-labelled in 14 programmable categories (type, color, segment thickness, etc.) and once positioned on screen have their characteristics in time and frequency transferred to any spreadsheet program such as Excel in one single mouse click (Fig. 1).

The user-defined categories can be adapted to virtually any phonological model of sentence intonation, using for instance either the ToBI or melodic contours notation.

In the incremental prosodic structure model for example (Martin, 2018), sentence intonation uses melodic contours categories C0 (terminal falling and low for declarative sentences), C1 (rising above the glissando threshold), C2 (falling above the glissando threshold) and Cn (rising or falling below the glissando threshold). The corresponding phonological category is automatically assigned by the software to the segment placed by the user, together with the color coding eventually assigned beforehand. Using ToBI notation such as H*H%, L*L-, , L+H* for FToBI, it is equally easy to predefine as actual acoustic melodic movements are then approximated in terms of tone targets. Labelling can then be done manually or automatically according to the user criteria. The integrated system has been applied to the analysis of a large number of speech archives, both old and recent. Examples pertaining to the changes occurred in the realization of melodic contours in French in the period 1913-2018 will be given, showing the capability of the system in degraded recording conditions to illustrate the evolution of prosodic style in political and news speeches of these periods.

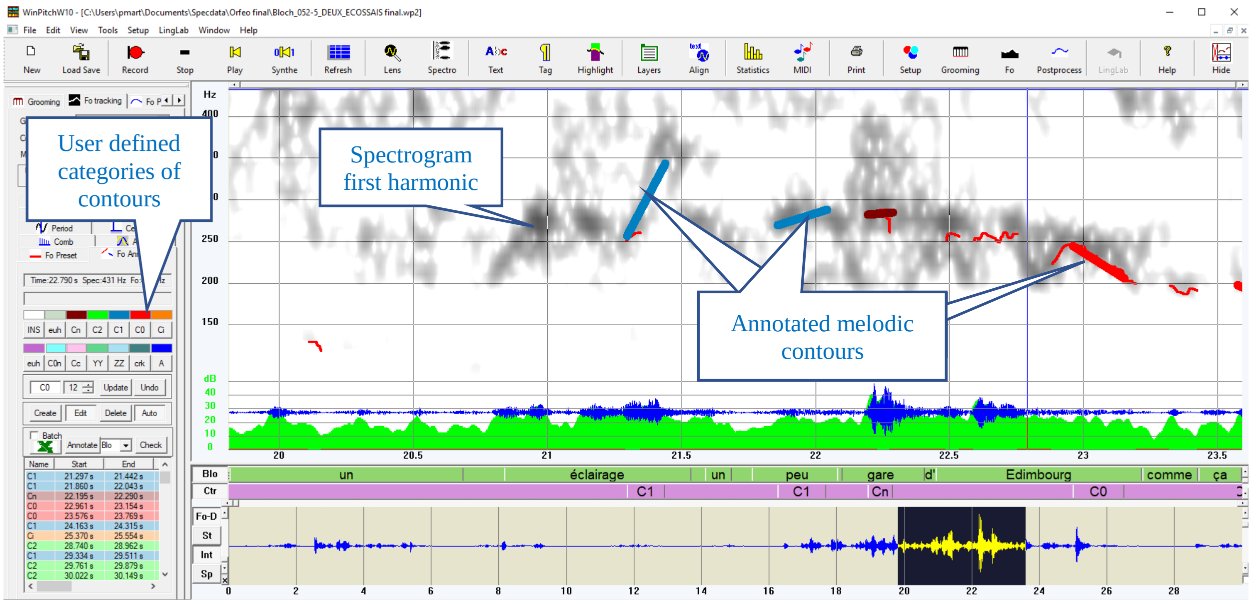

Fig. 1. An example of prosodic annotation of a noisy recording. Segments of the melodic contours are graphically annotated to fit both the pitch curve and the corresponding spectrogram first harmonic

Martin, Ph. (2018) *Intonation, structure prosodique et ondes cérébrales*, London : ISTE, 332 p.
WinPitch: www.winpitch.com