**Manual of the Praat Script**

***Beat Extractor***

***(*and *BeatExtractorSeveral)***

version 2003

by Plinio A. Barbosa (2003)

This manual gives a general overview of how the Beat Extractor Praat script works. The Beat ExtractorSeveral scripts does the same for several audio files with a particular extension (e.g., WAV).

Credit: Fred Cummins, for comments/suggestions in previous versions of this readme file.

This Praat script implements Fred Cummins' Beat Extractor (Cummins, F. and Port, R.,1998. J. Phon, 26(2), 145-171), with slight modifications. That, in turn, was developed from Sophie Scott's P-centre model (PhD thesis, UCL, 1994). The filter' bandwidth and shape are, however, different from that in Scott's work.

The present script generates a TextGrid file containing boundaries close to the vowel onsets (note that this is not the same thing as finding p-centres, which is an unsolved problem). Model parameters and default values were optimised using Brazilian Portuguese utterances. In the following a brief overview of what the programme does is presented. This may help in modifying the parameters for other languages or datasets (I also give Fred's default values, which were used with Irish English, as well as Scott's default values, which were used with British English).

**Input:** Any audio file with extension compatible with Praat.

**Output:** A TextGrid file with vowel onset to vowel onset intervals.

When running the script, three buttons allow: (1) to choose between a male and a female speaker (this option automatically chooses appropriate cut-off frequencies for the filter in step 1 below), (2) to choose between a Butterworth or a Hanning filter (step 1), and (3) to choose a technique for detecting boundaries, as explained in steps 4a and 4b below.

Steps (BeatExtractorFlux.doc gives the same information as a flow graph):

1. The speech signal is filtered with either a by-default second-order Butterworth filter, or a Hanning filter. This order of the first filter can be varied, but a filter with sharp skirts is not recommended (the order changing is my addition). The default order is 2 provided the value for the variable filter order 0(= auto) is not modified. The default cut-off frequencies for the Butterworth filter are 1000 Hz and 1800 Hz (the latter allows the detection of front vowels) for male speakers and (1150 Hz, 2100 Hz) for female speakers, assigned automatically, provided their values 0 (= auto) are not modified. This frequency band preserves F1 for low vowels and F2 for the others (since the filter skirts are relatively shallow, high front vowels are included). Scott used a Gamma tone filter with a center frequency of 597 Hz, and a band from 288 Hz to 909 Hz, approximately (but her interest was finding p-centres).

2. The filtered signal is rectified.

3. The rectified signal is low-pass filtered (variable Smoothing\_cut\_freq in the Praat form). A 20 Hz cut-off frequency is used as default (for technique 4a. 40 Hz is chosen instead, automatically, in technique 4b) , since in Brazilian Portuguese, fast intensity changes are produced with tap in intervocalic position, when both vowels are reduced (e.g., "xícara", cup). (Fred used 10 Hz, instead. Scott used 25 Hz).

This Praat script introduces another possible technique (compared to Fred's) for identifying specific points associated with local rises in amplitude, by first identifying those points where the rate of increase of the amplitude envelope is maximal:

4 (a). A beat is associated with a local maximum of the first derivative of the amplitude envelope (obtained after step 3), provided this maximum is higher than threshold 2 (expressed as a proportion of the maximum signal derivative amplitude. Default = 0.12) and the absolute value of the amplitude peak is higher than threshold 1 (default = 0.15) of the maximum signal amplitude (this constraint allows the algorithm to ignore steep onsets associated with very small rises in amplitude).

Cummins' original technique is also available with the following technique:

4 (b). A beat is associated with a local rise in the amplitude envelope of the signal obtained after step 3. I suggest using the point at which threshold 1 (default = 0.15) of the rise is complete. Cummins used the 0.5 point.

5. A TextGrid is generated with the above boundaries.

Thanks to Fred Cummins for the original beat extractor.