**Manual of the Praat Script**

***Prosody Descriptor Extractor***

version May 2020

by Plinio A. Barbosa (2014, 2020)

This manual gives a general overview of how the Prosody Descriptor Extractor Praat script works. Any questions can be asked directly to this email: pabarbosa.unicampbr@gmail.com. The script is documented directly thoroughly in each main section of the script file. You can have some info of what each part does by just scrolling down in the script text itself.

PART A – HOW THE SCRIPTS WORKS

Main operations:

1. The script starts with the input parameters form, the window that is presented after clicking on **Run**. The window is presented with the default values for the input parameters. The script, the reference file \*.TableOfReal which accompanies the script kit and all coupled Audio/TG files should be in the same folder.

For each parameter you may keep the value as it is informed (default) or writing another one if you want to modify the current value. They are:

Output file names. OutPutProsParameters, OutPutSil, OutPutEff and OutPutTones (in the script they correspond to the variables FileOutPar, FileOutSil , FileOutEff and FileOutTones respectively). Inform the complete name (with the extension TXT) of the output file, at your choice. Only OutPutProsParameters file will be generated in any case and from computations made in the labelled intervals of the obligatory chunk tier. OutPutSil file is written only if you have a Pause Tier; OutPutEff file is written only if you have a Vowel Tier and OutPutTones file is written only if you have a Tones Tier (see section B).

Audio files extension (AudiofileExtension variable). Inform the extension of the audio files to be analized preceded by “\*.”

Optional tiers presence. Inform, by checking the corresponding options, which are the optional tiers present in your TG file (variables HasTonesTier, HasVVTier, HasVowelTier, HasSilTier).

F0 unit. Inform, by checking (in semitones) or unchecking (Hertz) option InSemitones, which unit for F0 computation is desired. Semitones are computed using 1 Hz as reference.

Tiers' positions (tonesTier, vVTier, vowelTier, silTier and chunkTier variables). Inform the positions where the respective tiers are in your TG file.

Pitch cut-offs (F0Thresholdleft and F0Thresholdright variables). The default values (75 and 300 Hz) are for male speakers. This is the range where the Praat F0 tracker will search for F0 values. For very expressive speech you’ll probably have to use a higher upper value. For females you should change this to the range 120 to 500 Hz. If you know the F0 range of your data, inform the limits around 10% less for the lower limit and around 10 % more for the upper limit. If you have both female and male speakers, use from 75 and 600 Hz.

Smoothing degree for F0 smoothing (Smthf0Thr variable). The default value is 2 Hz. This is to compute F0 main peaks throughout the utterances. Tests can be done beforehand with the utterances to evaluate if this is the best solution for the language.

F0 first derivative step (F0step variable). It is the temporal step, in seconds, to be used to compute F0 first derivative. We recommend do not change the default value.

Spectral emphasis threshold (spectralemphasisthreshold variable). The default value is 400 Hz for a L0 fixed band for computing spectral emphasis in Traunmüller & Eriksson (1995) terms.

Reference language (Reference variable). Indicates the TableOfReal file corresponding to the language whose reference values for phone duration (mean and standard-deviation) need to be used to normalize raw syllable-sized duration. It is an accompany file for the script (BP.TableOfReal for Brazilian Portuguese. Please, ask the authors for tables for other languages).

1. The script continues by recovering the names of all wav files in the current folder. Then the algorithm runs for each audio file in the folder. Each paired-by-name TG file is also read.
2. The script computes obligatorily all parameters listed in item 5 for each chunk of the chunk tier.
3. Durations of VV units are normalized and smoothed according to segmentation in the VV tier, if any. In this phase, three output text files are produced with (a) the sequence of VV duration z-scores for each audio file as ***audiofilename*dur.txt**, (b) stress group duration and number of VVs in each stress group as ***audiofilename*SG.txt** and (c) a modified TextGrid file including an interval tier with the stress group intervals and a point tier with smoothed z-scores peak values at the end of the stress group, obtained from the SG Detector algorithm (Barbosa, 2006, 2007) as ***audiofilename*Enriched.TextGrid**. All VV units should be segmented and labelled according to ASCII symbols in TabelaIPAMarcacaoSGDetector.pdf, attached to the script kit. The two most common errors are: (1) error of labelling (in that case the script gives an error informing the wrong label and where in time its is located), (b) let a blank interval in any place from the first to the last VV (in that case the script gives an 'empty formula' error).
4. The parameters computed by the script from the **chunk tier** are (variable name in italics):
   1. F0 median in semitones re 1 Hz/Hertz assigned to the variable *f0med* in the output file,
   2. F0 standard-deviation in semitones/Hertz assigned to the variable *f0sd* in the output file),
   3. F0 semi-amplitude between quartiles in semitones/Hertz assigned to the variable *fSAQ* in the output file),
   4. F0 minimum in semitones re 1 Hz/Hertz assigned to the variable *f0min* in the output file,
   5. F0 maximum in semitones re 1 Hz/Hertz for the whole chunk assigned to the variable *f0max* in the output file,
   6. Standard-deviation of F0 maxima in semitones/Hertz assigned to the variable *sdf0peak* in the output file,
   7. Mean peakness of F0 max in semitones re 1 Hz/Hertz relatively F0 range and multiplied by 1000 assigned to the variable *f0peakwidth*  in the output file,
   8. Smoothed F0 peak rate in peaks per second assigned to the variable *f0peak\_rate* in the output file,
   9. Standard-deviation of the F0 maxima positions, in seconds, assigned to the variable *sdtf0peak* in the output file,
   10. 1st-derivative F0 mean in Hertz/frame of the positive derivatives assigned to the variable *df0posmean* in the output file,
   11. 1st-derivative F0 mean in Hertz/frame of the negative derivatives assigned to the variable *df0negmean* in the output file,
   12. 1st-derivative F0 standard-deviation in Hertz/frame of the positive derivatives assigned to the variable *df0sdpos* in the output file,
   13. 1st-derivative F0 standard-deviation in Hertz/frame of the negative derivatives assigned to the variable *df0sdneg* in the output file,
   14. Spectral emphasis in dB assigned to the variable *emph* in the output file,
   15. Intensity variation coefficient assigned to the variable *cvint* (standard deviation of global intensity divided by global intensity mean) in the output file,
   16. LTAS slope between bands 0-1000 Hz and 1000/4000 Hz assigned to the variable *slLTASmed* in the output file,
   17. LTAS slope between bands 0-1000 Hz and 4000/8000 Hz assigned to the variable *slLTAShigh* in the output file,
   18. HNR (Harmonic-to-Noise ratio) in dB assigned to the variable *hnr* in the output file,
   19. SPI (Soft Phonation Index) in dB assigned to the variable *SPI* in the output file. It is the average ratio of low frequency harmonic energy in the band 70–1600 Hz to the higher frequency harmonic energy in the band 1600–4500 Hz for the voiced areas of the analyzed signal. The higher the value the less adducted are the vocal folds during vibration,
   20. Local shimmer in percentual values assigned to the variable *shimmer* in the output file,
   21. Local jitter in percentual values assigned to the variable *jitter* in the output file,
   22. Speech rate in VV units per second assigned to the variable *srate* in the output file IF there is a VV tier,
   23. Articulation rate in VV units per second assigned to the variable *artrate* in the output file IF there is a VV tier and a Sil Tier.
5. The parameters computed by the script from the **tones tie**r are (variable name in italics):
   1. Tone types along the tier, as labelled by the user, assigned to the variable *tonetype* in the tones output file,
   2. Time where the tone is in the audio file in seconds, assigned to the variable *time* in the tones output file,
   3. Relative alignment from to VV unit onset divided by the VV duration, assigned to the variable *alignVV* in the tones output file,
   4. Mean F0 of the VV unit where the tone is, assigned to the variable *meanf0VV* in the tones output file,
6. The parameters computed by the script from the **sil tie**r are (variable name in italics):
   1. Pause type in the intervals, as labelled by the user, assigned to the variable *type* in the sil output file,
   2. IPI, the duration from the onset of the previous pause to the onset of the current pause interval, assigned to the variable *IPI* in the sil output file,
   3. Pause duration, the duration of the pause interval, assigned to the variable *durSIL* in the sil output file.
7. The parameters computed by the script from the **vowel tie**r are (variable name in italics):
   1. Vowel segment, the label of the vowel of each interval, assigned to the variable *vowel* in the vowel output file,
   2. H1- H2, the value in dB of the difference between harmonics H1 and H2 in the central part of the vowel interval, assigned to the variable *H1H2* in the vowel output file,
   3. CPP, the value in dB of the Cepstral Prominence Peak in the central part of the vowel interval, assigned to the variable *CPP* in the vowel output file.
8. The **VV tie**r is used to generate the files designated in item 4 above.

PART B – HOW ANNOTATION SHOULD BE DONE

Annotation should be done with Praat TextGrids with a mandatory chunk tier containing the intervals within which to compute the prosodic parameters described above for its tier. The symbols for the phonetic segments in the VV Tier must coincide with those indicated in the language-specific TableOfReal file. Adaptations can be done under request to P. Barbosa. The rest is done automatically by the script.

The figure below illustrates all tiers. Their names appear in the right part of the panel. We can see that, in the Sil Tier, there are silent pauses (#) and a filled pause (i:). When segmenting these pauses, it is important to make sure that the temporal limits of the corresponding chunk tier includes all its pause intervals of the Sil Tier. In the vowel tier there is only an interval with vowel /a/, but it can have much more, any vowel from which to extract H1-H2 and CPP values. The VV Tier is the usual VV tier used for normalizing duration in the SG detector script (2006): it is the sequences of all V-V intervals of the audio file with no blank intervals in between. Labels should be in accordance with the symbols in the attached *TableIPA-SGDetectorSimbols.pdf* file. The Chunk Tier indicates with two labels the intervals within which the prosodic parameters in item 5 above will be computed (in the example here, trecho1 and trecho2). The Tones Tier has only two tones here and it is a point tier.

