

Intensity_mean_3_0.praat

Task:

This script opens all sound files in a directory and associated TextGrids (if they exist), computes intensity-values at the center and/or edges of all segments, which can be specified in various ways, together with means and quantiles and writes the results to a text file with the name “intensity_results_<date>_<time>.txt”. The starting position and the length of the segment are reported along with the intensity data. Sound files without an associated TextGrid file will be treated as one segment on its own.

There are more parameters that control the behaviour of the script than displayed in the form window (the size of the form window fits into a screen with 900+ points vertical resolution). These parameters can be set below the form statement in the script (see the **Programming** section).

Parameters:

The screenshot shows a dialog box titled "Run script: Intensity parameters:" with the following fields and callouts:

- Directory path:** A text input field at the top, with a callout pointing to it.
- Tier number to be analysed:** A text input field containing the number "1", with a callout pointing to it.
- Specification of labels:** A text input field containing a period ".", with a callout pointing to it. Above it is the text "<label>, <list>.txt, '!' (= only labelled), or empty (= all)".
- Type of intensity measure:** Three radio buttons labeled "energy", "sone", and "dB". The "dB" button is selected, with a callout pointing to it.
- Report data from a center of a segment:** A checkbox labeled "Center position", which is unchecked, with a callout pointing to it.
- Report data from the left and right edges:** A checkbox labeled "Edges", which is unchecked, with a callout pointing to it.
- Report means and st. dev. of a segment:** A checkbox labeled "Means", which is checked, with a callout pointing to it.
- Specification of quantiles to be reported:** A text input field labeled "Quantiles:" containing an empty field, with a callout pointing to it.
- Reporting of skipped intervals:** Two radio buttons labeled "All" and "None". The "None" button is selected, with a callout pointing to it.
- Symbol to be used for missing data values:** A text input field labeled "Missing values symbol:" containing a period ".", with a callout pointing to it.

At the bottom of the dialog box are four buttons: "Standards", "Cancel", "Apply", and "OK".

Please read the **Programming** section for more internal parameters that can easily be changed.

Directory path:

The script handles all sound files in a directory specified in this field. If this field is left empty, the script will handle all sound files in the directory where the script was started (i.e., the script is placed in the same directory as the sound and TextGrid files).

Tier number to be analysed:

The number of the interval tier to be analysed.

Specification of labels:

The segments (= intervals in PRAAT terminology) that should be analysed can be specified in several ways (in case no TextGrid file is found, this field is ignored):

<label>:

Giving a label (e.g. a:) or a list of labels separated by commas or spaces (e.g. i:,I u:,U) will only report segments that have this label. This function is case sensitive.

<list>.txt:

Giving a text file (e.g. label_list.txt) will report all segments that are listed on a line-by-line basis in a raw text file (not a Word or Pages file). Note that the extension .txt must be given in this field. Example of such a text file:

```
a
a:
ae
0
```

‘.’:

Using a dot (.) will report values for every labelled segment.

empty:

Leaving this field empty will report values for all labelled and unlabelled segments.

Report data from a center of a segment:

The middle of each segment specified by **Specification of labels** will be reported. If no TextGrid file is found, the center of a sound file will be used.

Report data from the left and right edges:

The edges each segment specified by **Specification of labels** will be reported. If no TextGrid file is found, the center of a sound file will be used.

Report means and st. dev. of a segment:

Setting this switch will report the mean, standard deviation and median of each segment selected by **Specification of labels**. If no TextGrid file is found, the whole sound file will be used.

Specification of quantiles to be reported:

Quantiles of a segment will be reported by specifying values separated by commas or spaces. Values can be given between 0 and 1 (as fraction of a the whole range) or as percentages between 1 and 100. The program will always makes the reporting symmetrical to the lower and upper quantile range. Specifying quantiles will always force the median to be reported. For example, a specification like 0.05,10 will report the 5%, 10%, 50%, 90% and 95% quantiles. If no TextGrid file is found, the whole sound file will be used.

Reporting of skipped intervals:

Handling of segments which are excluded by the **Specification of labels:**

All:

All segments excluded from the computations are reported with File, Label, Start(s) and Duration(ms) only (see **Result file** below), all other values are set to the missing value symbol (see below). This function can be helpful to see the context of a particular segment analysed.

None:

No segments that are excluded are reported.

Symbol for missing values:

When PRAAT cannot compute a value it uses internally the string “- undefined -”. This script replaces this string by the string given in this field. For a subsequent analysis of the data with JMP, the dot indicate missing values, for R it would be NA.

Result file:

The script generates a raw text file with tab-delimited data and a header line. The file name is of the form “intensity_results_<date>_<time>.txt” with <date> of the form ‘yymmdd’ (i.e. 2-digit year, 2-digit month, 2-digit day) and <time> of the form ‘hhmmss’ (i.e. 2-digit hour in 24 hour format, 2-digit minutes, 2-digit seconds). For example, a file with the name “intensity_results_190322_150110.txt” was created on the 22nd of March 2019 at 3pm, 1 minute and 10 seconds. The parameters controlling the computation are listed at the end of every result file.

The result file for a parameter setting to report mean, standard deviation, median, 10% centils and values at the the left and right edges of an interval will look like:

File	Label	Start(s)	Duration(ms)	Mean (dB)	StDev (dB)	10% (dB)	Median (dB)	90% (dB)	Left (dB)	Right (dB)
g071a000	n	1.0455	82.5	71.56	1.53	68.84	71.89	72.64	72.68	68.70
g071a000	t	1.1281	14.5	69.69	.	70.33	70.33	70.33	68.71	70.37
g071a000	n	1.3933	78.6	72.19	1.63	69.37	72.54	73.23	72.18	66.37
• • •										
g071a000	t	8.2948	70.4	38.27	20.51	-4.08	25.35	44.63	-12.27	46.53
g071a000	n	8.8649	77.6	67.25	0.12	67.07	67.25	67.36	66.99	67.12

Analysis started: 30-Sep-20 14:12:21

Tier: 4

Labels: list.txt

Step rate: 0 s

Low FO: 50 Hz

Computation units: energy

Information for programming:

Some parameters can be set underneath the ‘form’ section in the script. These are:

Directories:

The script uses internally separate strings for sound, TextGrid, result and support directories. Users who use separate directories for these can adjust these names in the script (or put them into PRAAT’s ‘form’ window).

Sound file names:

The default extension for sound files is “.wav”. This parameter can be changed in the script.

User feedback:

The script reports which file is being handled and the percentage of all files in a directory that have been handled. By setting this switch, any output (other than error and warning messages) will be suppressed. This will decrease processing time, but there is no feedback other than an increasing size of the result file.

user_feedback = 0 Gives user feedback

user_feedback = 1 No user feedback

Noprogress string:

PRAAT itself reports its activity when computing pitch, intensity and formants. All these outputs can take substantial processing time, sometimes longer than the actual computation time. (This string is actually positioned where in PRAAT the `noprogress` is written.)

np_string\$ = "" (= empty string) normal PRAAT feedback

np_string\$ = “noprogress” No PRAAT feedback

Dummy data header line:

Statistic programs like JMP decide the type of data for each column on basis of the first data line. To force correct data-type assignment (due to missing data in the first data row) a dummy data line of text, 0 and 0.0 can be generated to force correct data-type assignment.

dummy_data_header = 0 No dummy data line

dummy_data_header = 1 Dummy data line with “Dummy” for strings and “0.0” for numerals

Duration reporting:

duration_in_ms = 1 Duration is reported in milliseconds

duration_in_ms = 0 Duration is reported in seconds

Intensity computing parameters:

step_rate = 0

low_F0 = 50

Current version and date:

3.0, 30-sep-2020

Known problems:

None

Planned extension:

Handling of point tiers

Reporting time points of measurements as an option

Contact:

reetz.phonetics@gmail.com

