

# Rater training / Rating reliably: Speaking



Nivja de Jong & Carol Spoettl



# Overview

- Rating analytically on fluency
- Measuring fluency
- Measuring fluency automatically



# Rating analytically: fluency

- Cognitive fluency, utterance fluency, and perceived fluency (Segalowitz, 2010)



# Aspects of fluency

- Cognitive fluency: ability of the L2 speaker to smoothly translate thoughts to L2 speech
- Perceived fluency: subjective measure of what listeners perceive – about L2 speaker's cognitive fluency
- Utterance fluency: objective acoustic measures of an utterance

# Utterance Fluency

Tavakoli & Skehan (2005)

- breakdown fluency
- speed fluency
- repair fluency

# To what extent do perceived and utterance fluency overlap?

- Recent study (Bosker et al., 2013): a lot.
- 84% of variance between rated variance could be explained by actual measures in utterance fluency
- (Naïve raters with specific instructions)

# Ratings:

- Fluency, taking into account:
  - Silent and filled pauses (both number and length)
  - Speech tempo
  - Repetitions and repairs
- Scale of 0 – 6

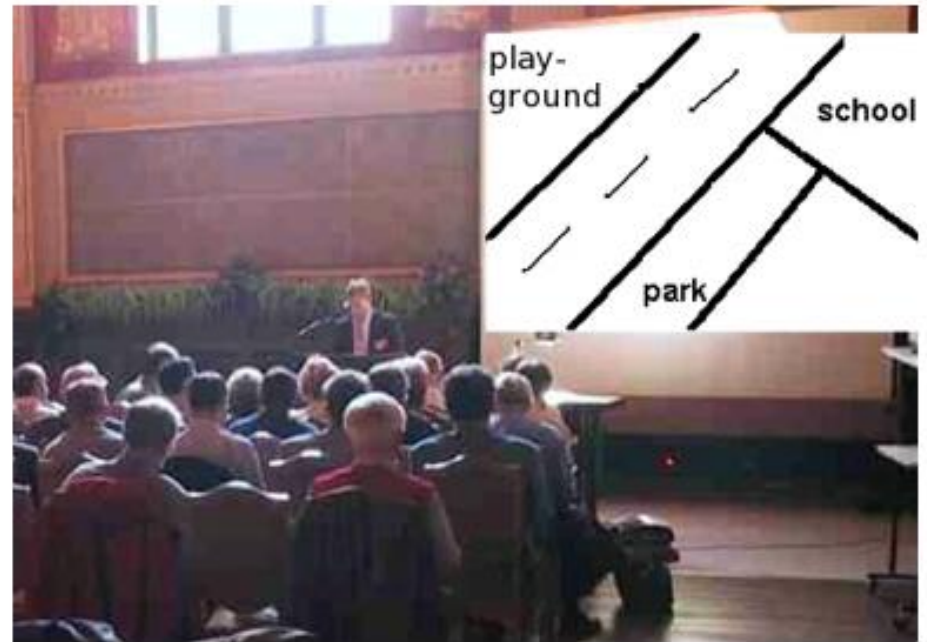


## Exercise 2: Comment in a Town Hall Meeting

You are attending a town meeting. Someone from the municipality (mister Smith), has presented a plan to build a new playground

(see picture to the right)

After the presentation, comments and questions are invited. You decide to express your opinion about the location and suggest an alternative.

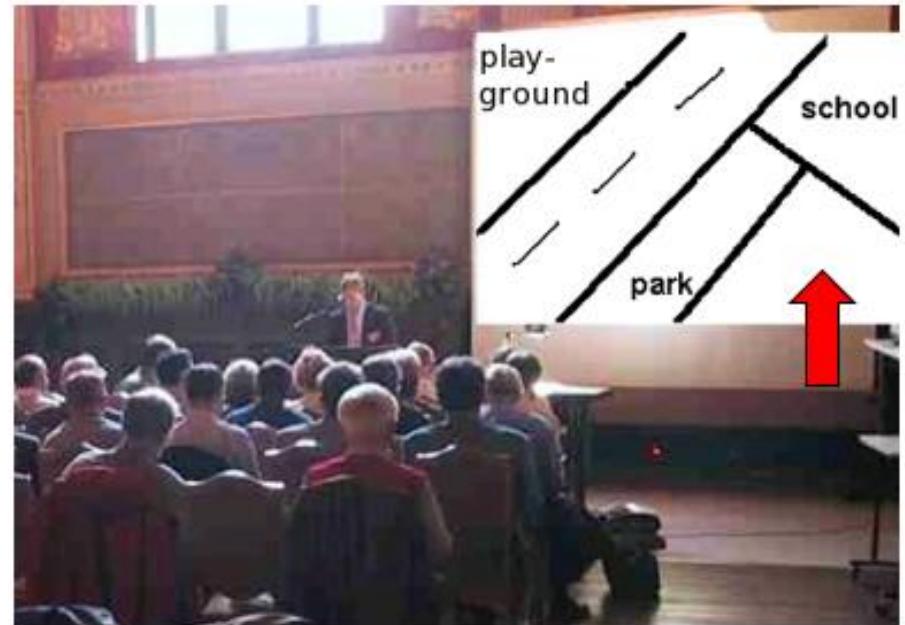


## Exercise 2: Comment in a Town Hall Meeting

You have raised your hand to get the floor. The arrow on the picture to the right shows your preferred alternative location.

Thank mister Smith (from the municipality) for his presentation

Explain the spot that you think could be a better location for building the playground and argue why this spot would be better.



Please prepare what you are going to say





# Ratings:


- Fluency, taking into account:
  - Silent and filled pauses (both number and length)
  - Speech tempo
  - Repetitions and repairs

- Scale of 0 – 6

Not fluent at all      0 1 2 3 4 5 6      very fluent

Pp29: 

Pp56: 

Pp70: 



# Measuring fluency

# Measuring Utterance Fluency

- Which measures?
- How to measure?

Tavakoli & Skehan (2005)

- breakdown fluency
- speed fluency
- repair fluency

# Choosing unconfounded measures

## Breakdown fluency:

Number of silent pauses / speaking time

Mean duration of silent pauses

Number of filled pauses / speaking time

Mean duration of filled pauses

## Speed fluency:

Articulation rate or mean syllable duration

## Repair fluency:

Number of repairs and or repetitions / speaking time

(Mean duration of repairs / repetitions)

# More on silent pauses

## Silent pauses

- Micropause (silence of .2s or less)
- Hesitation (silence of .3 to .4s)
- Unfilled pause (silence of .5s or more)

# Duration threshold of silent pauses

Occlusives such as “p”, “t”, “k”, etc: include short silences.  
These are ‘articulation’ pauses, and not ‘hesitation’ pauses.

Thresholds in literature between 100ms – 1000ms: ???

De Jong & Bosker (2013) advice to use a threshold of 250 – 300ms (higher correlation with overall proficiency measure)



# Measuring fluency automatically

# Automatic unconfounded measures

## Breakdown fluency:

Number of silent pauses / phonation time

Mean duration of silent pauses

Number of filled pauses / phonation time

Mean duration of filled pauses

## Speed fluency:

Articulation rate (or mean syllable duration)

## Repair fluency:

Number of repairs and or repetitions / phonation time

(Mean duration of repairs / repetitions)

# Automatic overall measure

Breakdown + Speed fluency:

Speech rate (number of syllables / total time)

# Measuring utterance fluency

- Record L2 speech-data
  - Annotate by hand (e.g.: PRAAT, CLAN)
  - For monologic speech (excerpts): use automatically obtained measures (De Jong & Wempe, 2009: script in PRAAT)
  - Use a combination of these two methods

# Praat script

- For each .wav-file in a directory, it will give you
  - nsyll
  - npause
  - dur (s)
  - phonationtime (s) (= speaking time)
  - **speechrate (nsyll/dur)**
  - articulation rate (nsyll / phonation time)
  - ASD (phonationtime/nsyll)

# Praat script HOWTO: download praat

- Go to [praat.org](http://praat.org) (or, search for: *praat*)
- Download “praat”
- Unzip
- Place “praat.exe” on your desktop  
(it is a small executable)



# Praat script HOWTO: get the script

- Open praat
- Go to <http://sites.google.com/site/speechrate>  
(or, search for: *speech rate praat script*)  
On the site, click on Praat script v2:  
as found in bottom left



Navigation
<a href="#">Speech Rate: Praat script that detects syllable nuclei</a>
<a href="#">How does the script work?</a>
▼ <a href="#">Praat Script Syllable Nuclei</a>
<a href="#">tutorial</a>
<a href="#">Praat Script Syllable Nuclei v2</a>

Copy the script:  
starting from “###” to “endfor” (bottom of page):

[Speech Rate: Praat script that detects syllable nuclei](#) >

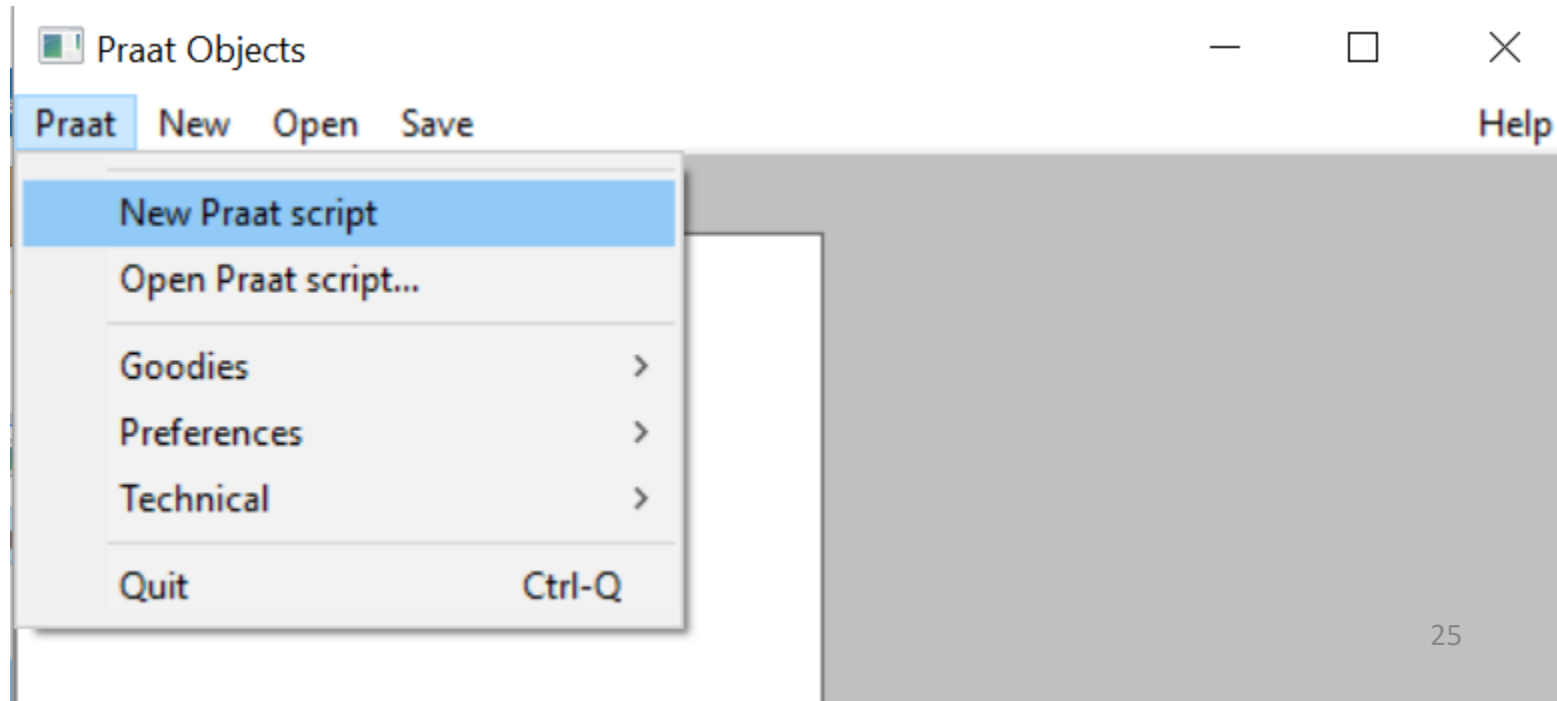
## Praat Script Syllable Nuclei v2

```
#####  
#  
# Praat Script Syllable Nuclei  
# Copyright (C) 2008 Nivja de Jong and Ton Wempe  
#  
# This program is free software: you can redistribute it and/or modify  
# it under the terms of the GNU General Public License as published by  
# the Free Software Foundation, either version 3 of the License, or  
# (at your option) any later version.  
#  
# This program is distributed in the hope that it will be useful,  
# but WITHOUT ANY WARRANTY; without even the implied warranty of  
# MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the  
# GNU General Public License for more details.  
#  
# You should have received a copy of the GNU General Public License  
# along with this program. If not, see http://www.gnu.org/licenses/  
#  
#####  
#  
# modified 2010.09.17 by Hugo Quené, Ingrid Persoon, & Nivja de Jong  
# Overview of changes:  
# + change threshold-calculator: rather than using median, use the almost maximum  
# minus 25dB. (25 dB is in line with the standard setting to detect silence  
# in the "To TextGrid (silences)" function.  
# Almost maximum (.99 quantile) is used rather than maximum to avoid using  
# irrelevant non-speech sound-bursts
```



# Praat script HOWTO: get the files

- Open PRAAT (double click)
- Ignore window “PRAAT picture” or close it
- In PRAAT Objects window, choose “New praat script”



# Praat script HOWTO: get the files

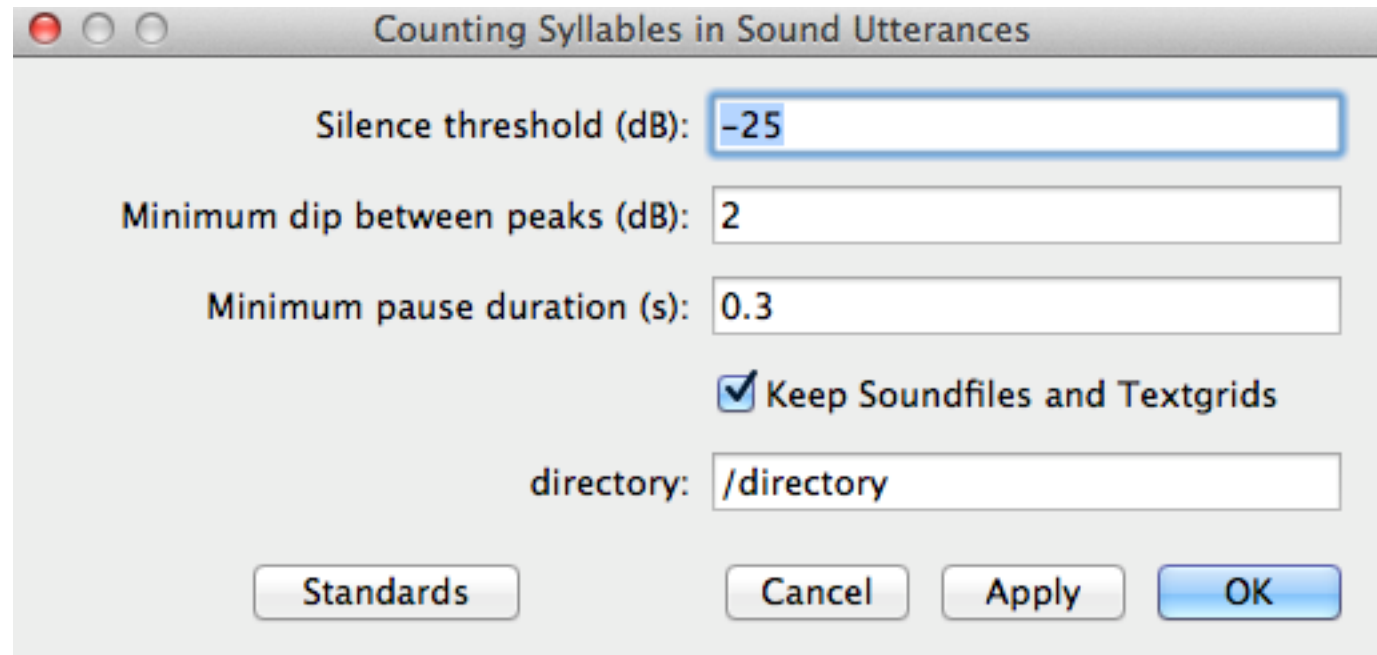
- Paste the script (ctrl-V) into the new window
  - NB on a mac, sometimes you need to paste it first into a text editor and then into PRAAT
- Download from the “summerschool”-page of the praat script some soundfiles, and save them into a folder that is easily accessible

<https://sites.google.com/site/speechrate/Home/summerschool>

# Praat script HOWTO: run the script

Run the script

- Ctrl-r or choose 'run script'

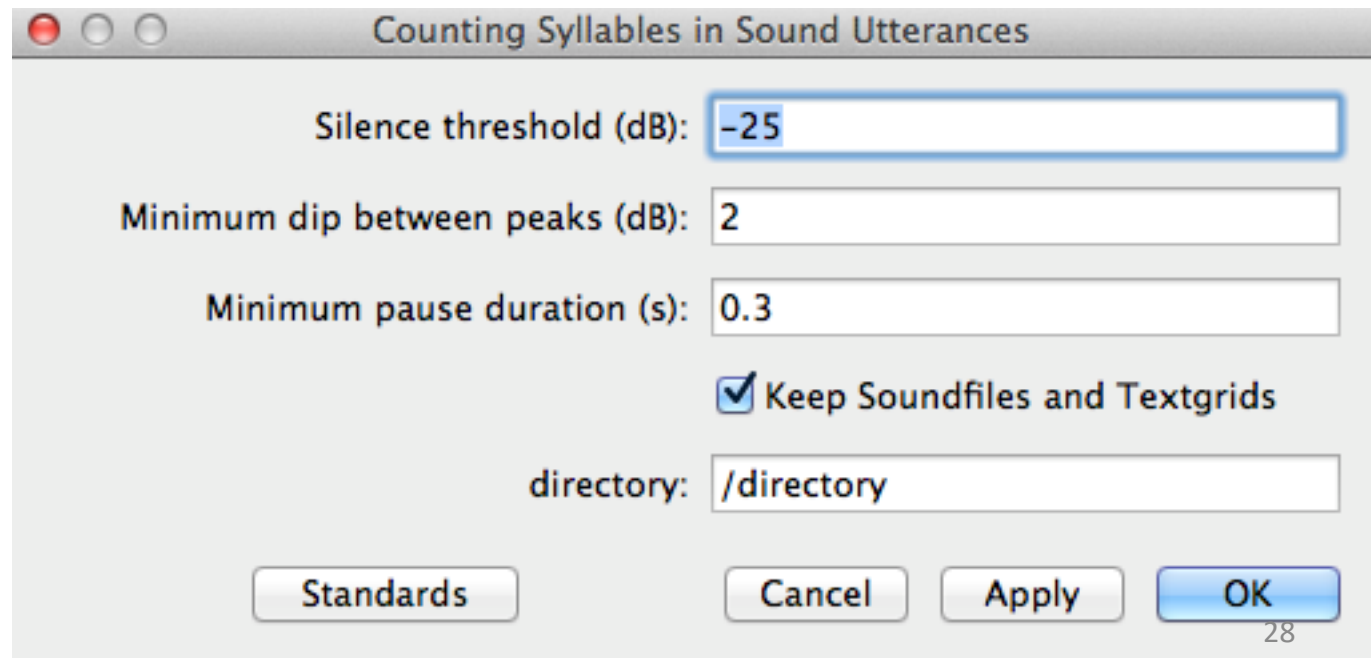


# Praat script HOWTO: change defaults

Change default values

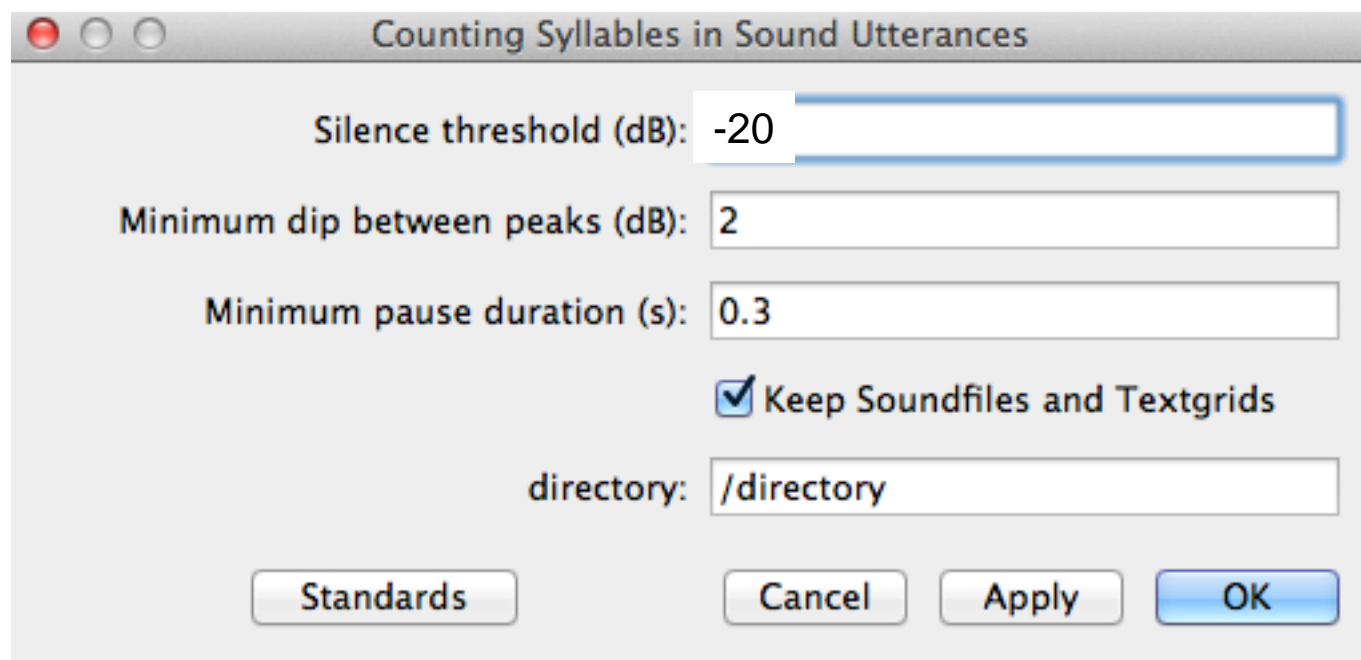
Change silence threshold “-25” becomes “-20”

Change “/directory” to your directory



Silence threshold (dB): from the 99% loudest peak, how much dB below should be counted as silence?

The higher this number (e.g. -30, -40), the lower the chance of finding silent pauses



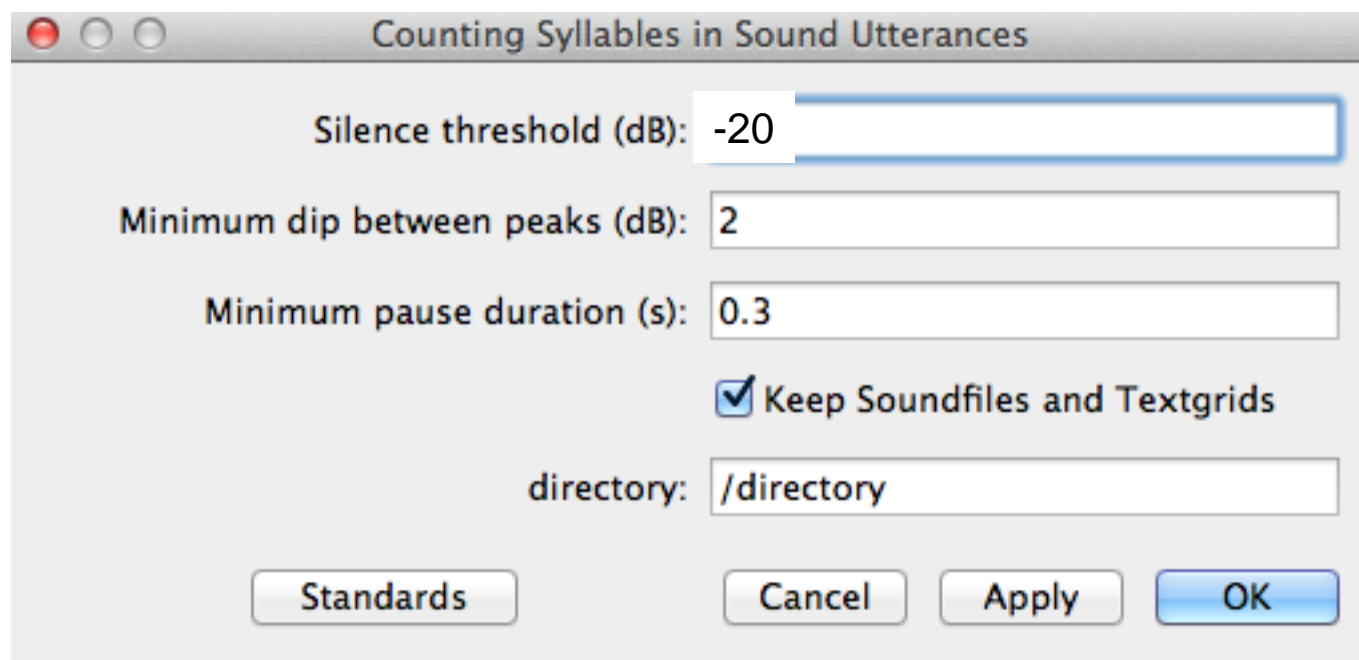
The image shows a screenshot of a software dialog box titled "Counting Syllables in Sound Utterances". The dialog box has a title bar with standard window controls (red, yellow, and green buttons). The main content area contains several input fields and a checkbox:

- "Silence threshold (dB):" with a text input field containing "-20".
- "Minimum dip between peaks (dB):" with a text input field containing "2".
- "Minimum pause duration (s):" with a text input field containing "0.3".
- A checked checkbox labeled "Keep Soundfiles and Textgrids".
- "directory:" with a text input field containing "/directory".

At the bottom of the dialog box, there are four buttons: "Standards", "Cancel", "Apply", and "OK". The "OK" button is highlighted in blue.

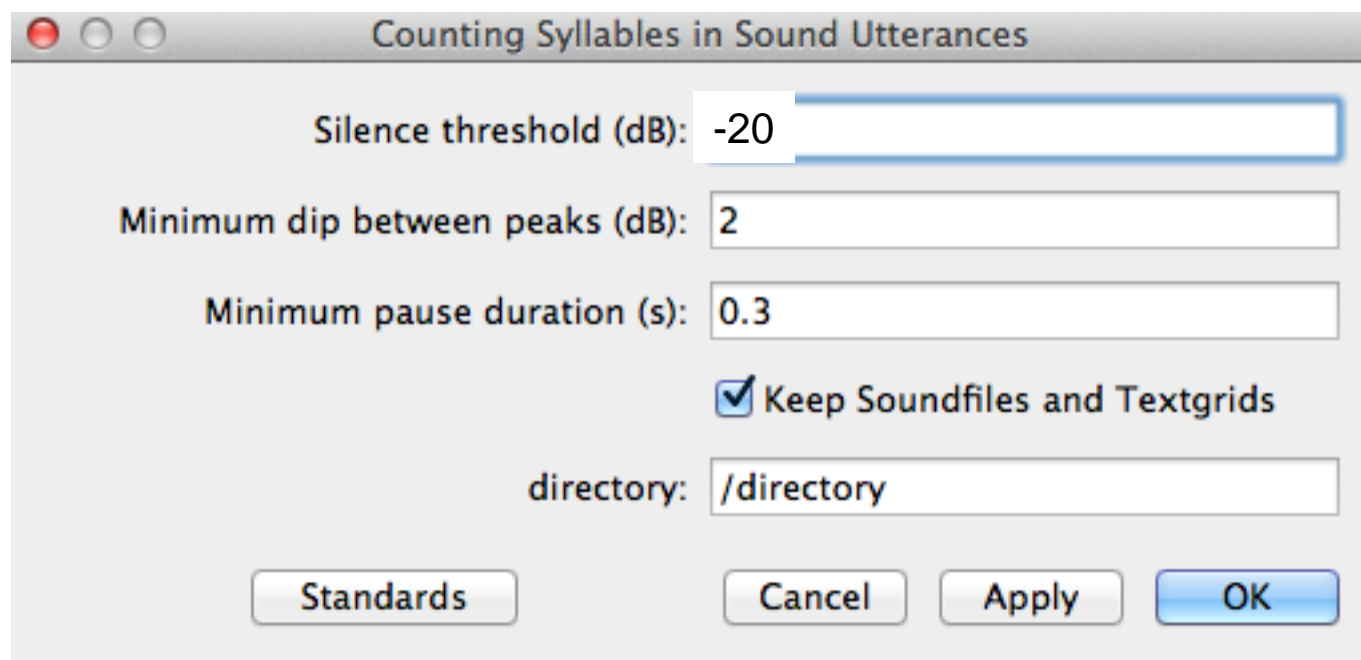
In between two consecutive peaks in loudness (“potential syllable nuclei”): how deep should the dip in intensity be?

The higher this number, the fewer syllables will be found



Minimum pause duration (s): How long should a pause be to be counted as a silent pause?

The higher this number, the fewer pauses will be found

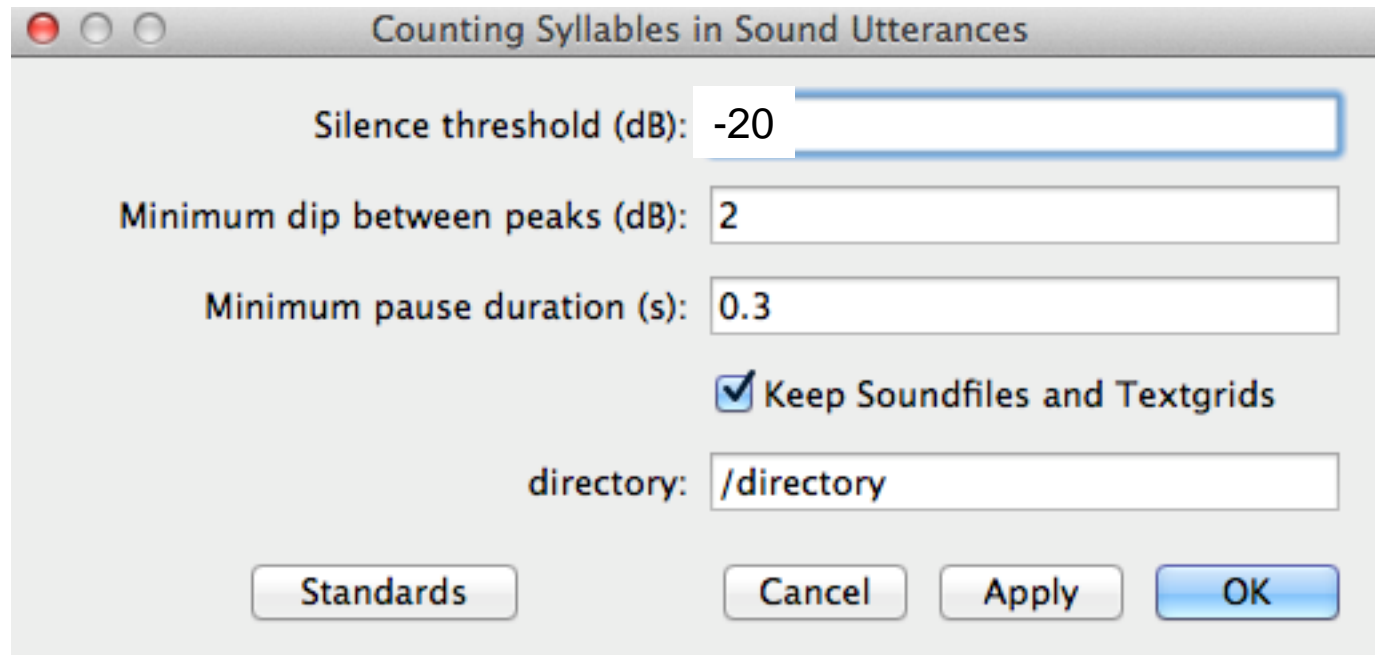


The image shows a dialog box titled "Counting Syllables in Sound Utterances". It contains several input fields and a checkbox. The "Silence threshold (dB)" field is set to -20. The "Minimum dip between peaks (dB)" field is set to 2. The "Minimum pause duration (s)" field is set to 0.3. There is a checked checkbox labeled "Keep Soundfiles and Textgrids". The "directory:" field is set to /directory. At the bottom, there are four buttons: "Standards", "Cancel", "Apply", and "OK".

Parameter	Value
Silence threshold (dB)	-20
Minimum dip between peaks (dB)	2
Minimum pause duration (s)	0.3
Keep Soundfiles and Textgrids	<input checked="" type="checkbox"/>
directory:	/directory

# Praat script HOWTO

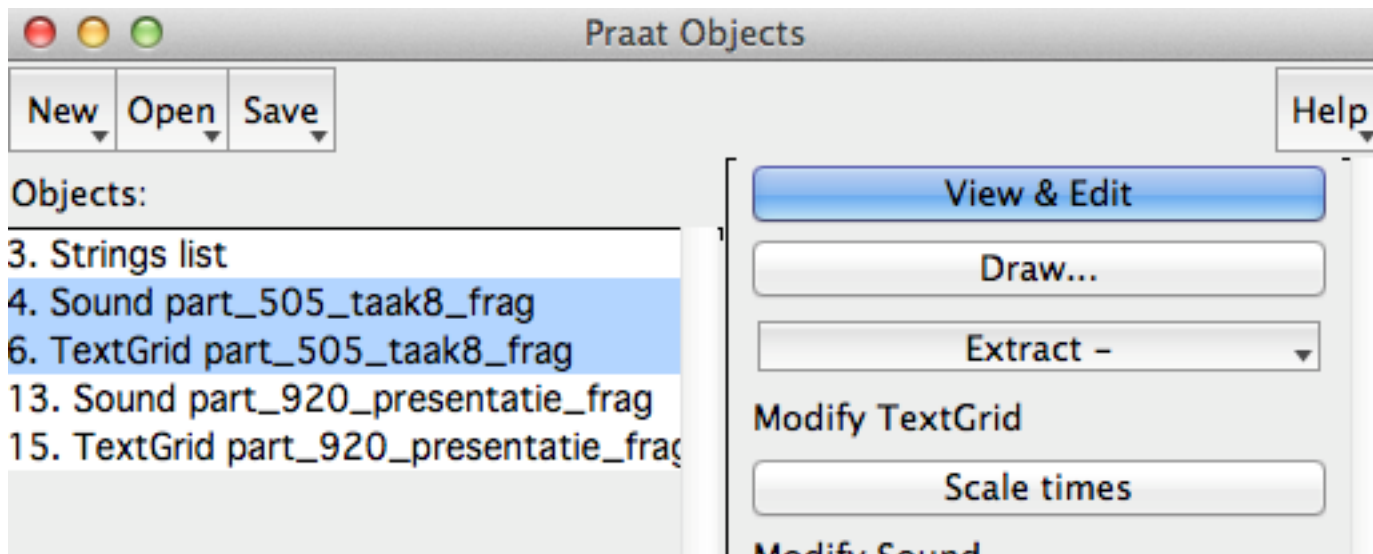
Check “Keep Soundfiles and TextGrids” to inspect performance of the script.





# Praat script HOWTO

Checking the TextGrids for performance:



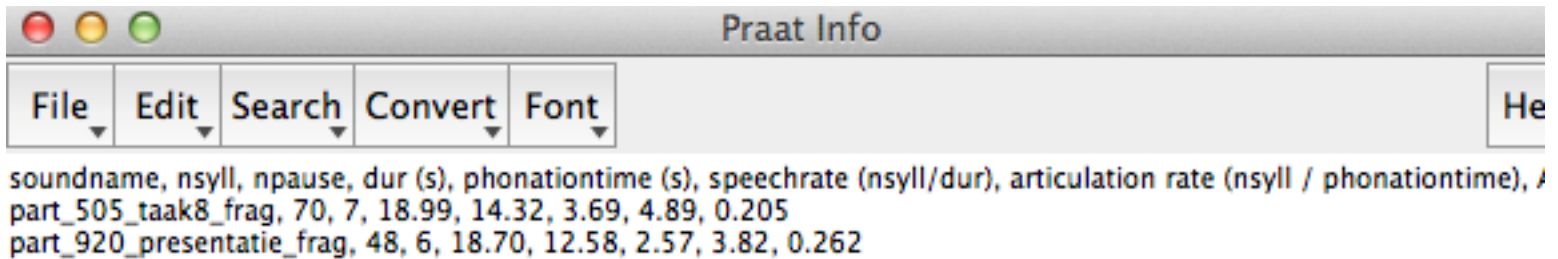
In the PRAAT objects window, select the Soundfile and the corresponding TextGrid together, then click “View and Edit”

# Praat script 'disclaimer'

- The script will only work if the soundfile does not have too much background noise (otherwise: perform a filter first)
- The script can only detect syllables that are actually there; many unstressed/reduced syllables not picked up
- A long syllable can erroneously be counted as two (if there are two peaks with a sufficient dip...)

# Praat script HOWTO

Results are written in an “Info”-window:



You can save this file, and open it in a spreadsheet program: indicate comma is a delimiter.

For now, just note the “speechrates” calculated by the script: 7<sup>th</sup> column (thus, 6<sup>th</sup> number)

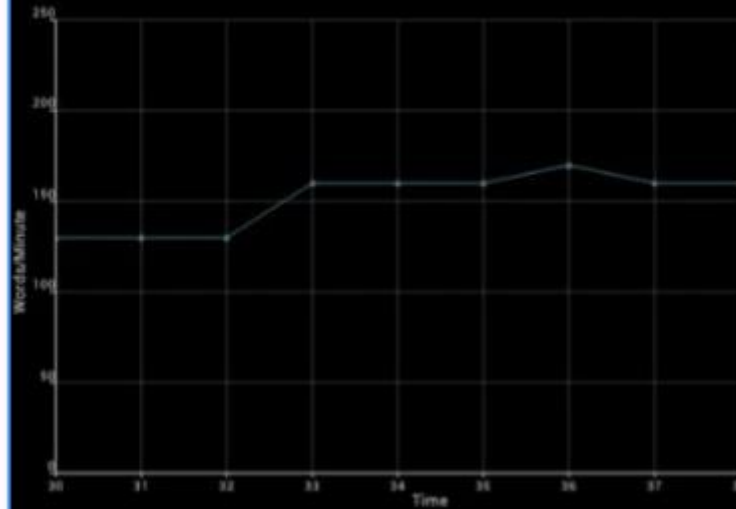
# Comparing ratings & measures

- Rank order the fluency ratings
- Rank order the speech rate measures
  
- Discuss discrepancies

PS funfact: App “speakrite” (available in Google Store) uses De Jong & Wempe (2009) algorithm



# SpeakRite Beta



# References

- Bosker, Hans Rutger, Anne-France Pinget, Hugo Quené, Ted J. M. Sanders and Nivja H. De Jong. 2013. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing* 30: 159–175.
- De Jong, Nivja H. 2016. Fluency in second language assessment. In Dina Tsagari & Jayanti Banerjee (Eds.) *Handbook of second language assessment*. (pp 203-218). Mouton De Gruyter.
- De Jong, Nivja H. and Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2): 385–390.
- De Jong, Nivja H. and Hans Rutger Bosker. 2013. Choosing a threshold for silent pauses to measure second language fluency. In: Robert Eklund (ed.). *Proceedings of Disfluency in Spontaneous Speech, DiSS 2013*, 17–20.
- Riggenbach, Heidi. 1991. Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes* 14(4): 423–441.
- Segalowitz, Norman. 2010. *Cognitive bases of second language fluency*. New York: Routledge.
- Tavakoli, Parvaneh and Peter Skehan. 2005. Strategic planning, task structure and performance testing. In: Rod Ellis (ed.). *Planning and task performance in a second language*, 239–276. Amsterdam: John Benjamins.