



Max Planck Institute
for Psycholinguistics

PRAATALIGN: PHONETIC ALIGNMENT MADE EASIER FOR PSYCHOLINGUISTIC DATA PROCESSING.

Will Schuerman¹, Mart Lubbers¹, & Francisco Torreira¹.
1. Max Planck Institute for Psycholinguistics

Manual Phonetic Segmentation

- Extremely time consuming and costly (1 min spontaneous speech == 13 hours of annotation) [1]
- Even when accompanied by phonetic transcriptions, manual forced alignment of a 15 min. corpus may require two hours of work. [2]

Praatalign: Forced Alignment in Praat [3]

The image shows three screenshots from the Praat software. The left screenshot shows a waveform and spectrogram of the sentence 'de aap is boven de raaf' with manual segmentation intervals. The middle screenshot shows the 'Praatalign' menu with options like 'Align interval', 'Add interval on tier 1-8', and 'Setup force alignment...'. The right screenshot shows the same audio with Praatalign applied, resulting in automatic segmentation intervals.

Psycholinguistic Experimentation

- Often uses tightly controlled stimuli in which the to-be-produced content is known beforehand.
- Unless it is the focus of the study, ignores phonetic variation in vocal responses (due to time/cost factors)
- Forced alignment may enable time- and cost-effective analysis of phonetic variation in such experiments.

Manual vs. Automatic Segmentation

- Data derived from recordings obtained for experiment on speech perception. [4] 28 Speakers x 120 Words, read in citation format. Phonemic transcriptions obtained from DutchPond Database [5].

Segment Onset Comparison	<20ms	<10ms
H1 vs. Praatalign	64%	45%
H1 vs. Praatalign (Voiced Stops)	89%	78%
H1 vs. H2 (From [2])	87%	57%

PraatAlign vs. To Silences...

- Speech Onset Latency of raw recordings obtained in [4].
- Praatalign takes into account average duration of different phones.

Average SOL: To Silences...	487.76(161.95)
Average SOL: Praatalign	445.63 (151.68)

The image shows two screenshots of the Praat software. The left screenshot shows manual segmentation of the word 'koor' into segments for '<', 'k', 'o:', 'r', and '#'. The right screenshot shows the same audio with Praatalign applied, resulting in segments for 'silent', 'sounding', and 'silent'.

Discussion

- Unsupervised Praatalign forced alignment performs reasonably well in a fraction of the time. Excellent accuracy for voiced stops.
- Semi-supervised alignment and correction of outliers is also possible via interactive Praat scripting, when detail is required (such as VOT measurement).

References

References: [1] Schiel & Draxler (2003), Munich; [2] Goldman (2011), Interspeech'11; [3] Praatalign: an interactive Praat plug-in for performing phonetic forced alignment. A detailed manual for version 1:3. http://dopefishh.github.io/praatalign/book/book_1.3.pdf; [4] Schuerman, McQueen, & Meyer (2015) [5] Marian, Bartolotti, Chabal, & Shook (2012), PLOS One

For More Information On Use and Installation, Please Visit

<http://dopefishh.github.io/praatalign/>

Will.Schuerman@mpi.nl