Using confidence intervals in forced alignment

Matthew C. Kelley, George Mason University, USA
mkelle21@gmu.edu

Forced alignment is a process by which a transcription is automatically aligned in time with a speech signal. In experimental phonetics, forced alignment tools automate the laborious task of creating time-aligned segment- and word-level transcriptions.

Most forced alignment systems calculate only a point estimate of each segment boundary. The time points in the segmentation yield an optimal alignment between the transcription and acoustics. However, in addition to the point estimate of the boundary, the Mason-Alberta Phonetic Segmenter (MAPS, Kelley et al., 2024) can output boundaries with 97.85% confidence intervals (Kelley, 2025). MAPS constructs intervals using order statistics on an ensemble of boundaries produced by several acoustic models. In addition to traditional word- and segment-tiers in a Praat TextGrid, MAPS also yields confidence intervals in a point tier. An example alignment with is given in Figure 1.

The present project aims to show how these confidence intervals on segment boundaries can be used fruitfully in phonetic research. Examples include detecting poor alignments, identifying segment sequences that are difficult to separate, and quantifying the system's uncertainty in the segmentation. In addition, the process of creating the confidence intervals will be detailed to open more avenues for methodological inquiry into generating confidence intervals for segmentation.

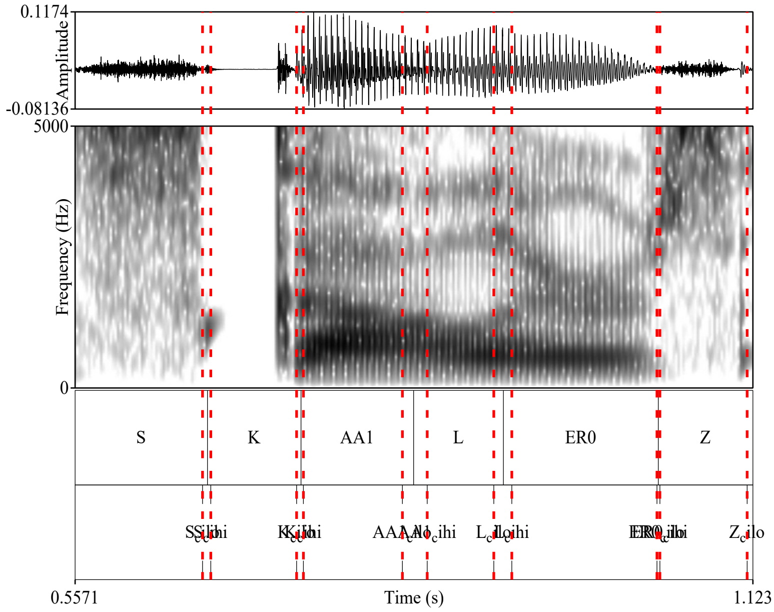

Figure 1. Sample CIs for scholars [ˈskalɚz] from ...blical scholars arg... from"biblical scholars argue" in TIMIT corpus. High and low ends of the CIs are drawn in dashed red lines. Note that the CIs are not symmetrical due to the order statistics method.

References

Kelley, M. C., Perry, S. J., & Tucker, B. V. (2024). The Mason-Alberta Phonetic Segmenter: A forced alignment system based on deep neural networks and interpolation. *Phonetica, 81*(5), 451–508. https://doi.org/10.1515/phon-2024-0015

Kelley, M. C. (2025). *Confidence intervals for forced alignment boundaries using model ensembles*. arXiv. https://doi.org/10.48550/arXiv.2506.01256