Utterance rhythm and tempo perception

Rachel Smith, Leendert Plug
University of Glasgow, University of Leeds
rachel.smith@glasgow.ac.uk, l.plug@leeds.ac.uk

Although research on rhythm perception and measurement increasingly highlights the complex relationship between rhythm and tempo (e.g. Arvaniti 2012, White et al. 2012), research on tempo perception and measurement has largely proceeded on the assumption that rhythm and tempo are separate dimensions of the temporal organisation of speech. Notably, there has been no research to date on the extent to which listeners take utterance rhythm — variation in utterance-internal temporal organisation — into consideration when estimating utterance tempo. Our study begins to fill this gap.

We follow up on an experiment reported in Plug & Smith (2017), in which listeners heard pairs of utterances of the general shape *this N or that N*, where 'N' stands for a noun, and had to decide for each pair how the utterances compared in tempo. We manipulated segment rate by varying the phonological complexity of the Ns, while keeping utterance duration constant. Among the findings was that listeners showed sensitivity to the relative duration of the final N: when relatively great, listeners heard the utterance as relatively fast. One explanation of this is that English listeners expect utterance-final lengthening, and therefore leave final syllables aside when making utterance tempo judgements. Another is that listeners pay crucial attention to function words when making tempo judgements, as these are particularly prone to temporal compression when tempo increases.

In the current experiment we used a similar design, but systematically varied the relative durations of the Ns and the function word constituents *this* and *or that.* The two utterances within each pair had identical segment rates and overall durations, but different Ns: e.g. *this tusk or that stick ~ this test or that step.* We manipulated constituent durations so that we could pair each first utterance with four rhythmically different versions of the second: (1) a naturally timed version, (2) one with the duration of *this* and *or that* expanded by 15% (and the Ns compressed to accommodate this), (3) one with the duration of the two Ns expanded by 15% (and *this* and *or that* compressed), and (4) one with the duration of only the second N expanded by 15% (and the rest of the utterance compressed).

Preliminary analysis shows that while expanding the durations of *this* and *or that* has little impact on listeners' tempo judgements, utterance versions with expanded Ns — and therefore compressed function words — are perceived as faster than naturally timed versions. The effect is strongest when both Ns are expanded, but already strong when only the second N is, consistent with Plug & Smith (2017). We will present a full analysis of these results, and relate them to previous findings on tempo and rhythm perception.

References

Arvaniti, A. 2012. The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics* 40, 351–373.

Plug, L. & R. Smith 2017. Phonological complexity, segment rate and speech tempo perception. *Proceedings of Interspeech*.

White, L. et al. 2012. Language categorization by adults is based on sensitivity to durational cues, not rhythmic class. *Journal of Memory and Language* 66, 665–679.