

Speech rate mediated compensation for assimilation in spoken word recognition

David Cheng-Huan Li & Elsi Kaiser (University of Southern California)

lidc@usc.edu

Spoken word recognition; Speech rate; Visual world paradigm

The role of speech rate has received relatively little attention in spoken word recognition, although speech production studies have shown that it can alter patterns of coarticulation and assimilation (e.g. [1]). Previous studies have shown that listeners can compensate for coarticulation, perceiving identical sounds differently in different contexts (e.g. [4]). In addition, listeners may also compensate for assimilation ([3]), a phenomenon observable in sentences like “A quick *run* picks you up.” Here, *run* may sound like *rum* when /n/ becomes assimilated to /m/ due to the influence of /p/. We wanted to test whether listeners’ processing of (potentially) assimilated sounds is sensitive to speech rate. Given that faster speech rates are associated with more coarticulation, *run* may be more acoustically similar to *rum* at faster speech rates than slower speech rates. Therefore, when identical lexically-ambiguous forms resulting from coronal assimilation are presented to listeners, we predict that these forms should be (i) more likely to be interpreted as the unassimilated form (e.g. *run*) when embedded in fast speech rates but (ii) more likely to be interpreted as the assimilated form (e.g. *rum*) when embedded in slower rates.

To test whether speech rate influences the interpretation of lexically-ambiguous forms, we conducted a **visual-world eye-tracking study**. Target sequences (noun+verb) were spliced into carrier phrases with different speech rates (fast/medium/slow, ex.1). The critical noun+verb sequences had two different coarticulation overlap patterns between the noun and verb (less overlap=more /n/-like vs. more overlap=more /m/-like), elicited with different prosodic contexts. Thus, we had a 3x2 design (36 targets, 50 fillers). To avoid neighborhood-density or plausibility confounds, critical nouns were nonwords: Participants were taught the nonwords were the names of aliens. Participants (n=24) listened to sentences and saw displays showing different aliens, two of which were the critical images (e.g. pictures of aliens called *Vone* and *Vome*). They were instructed to click on the alien mentioned.

Results: Eye-movement patterns show speech rate x overlap interaction 100-300ms and 400-600ms after alien-name onset ($p < .05$). More specifically, in the ‘more overlap’ conditions (more /m/-like), participants are more likely to look to the unassimilated form (e.g. *Vone*) in the fast and medium speech-rate conditions, but in the slow speech-rate condition, there are more looks to the assimilated form (e.g. *Vome*). In the ‘less overlap’ conditions (more /n/-like), participants tend to look at the unassimilated form in the slow speech-rate condition instead.

Conclusions: Even with identical acoustic materials, listeners’ online interpretations of ambiguous words are influenced by the speech rate of the surrounding carrier sentence. These findings are consistent with work on speech rate and perception of function words ([2]), as well as research showing that listeners are sensitive to sub-phonemic variation (e.g. [5]). While existing theories may accommodate our findings in different ways, the general point that emerges from our study is the importance of including speech rate in models of spoken recognition.

(1)	Carrier phrase part 1	Target N+V sequence	Carrier phrase part 2
	Every time the waiter brings out a strawberry cheesecake from the kitchen,	Vone peeks or Vome, peeks	to see if he can steal a piece.
	Fast, medium, or slow	Less or more overlap	Fast, medium, or slow

References

- [1] Browman & Goldstein. 1990. Tiers in articulatory phonology. *Papers in Laboratory Phonology 1*.
- [2] Dilley & Pitt. 2010. Altering context speech rate can cause words to (dis)appear. *Psychological Science*
- [3] Gaskell & Marslen-Wilson. 2001. Lexical ambiguity resolution and spoken word recognition. *JML*.
- [4] Mann. 1980. Influence of preceding liquid on stop-consonant perception.. *Perception & Psychophysics*
- [5] Salverda et al. 2007. Effects of prosodically modulated sub-phonetic variation on lexical competition. *Cognition*