

Integrating prosodic and discourse cues in spoken-word recognition

Meredith Brown, Anne Pier Salverda, Christine Gunlogson, & Michael K. Tanenhaus
University of Rochester
mbrown@bcs.rochester.edu

Listeners interpret segmental lengthening as evidence for an upcoming prosodic boundary, favoring lexical candidates whose word boundaries align with the perceived prosodic boundary (Salverda, Dahan, and McQueen, 2003). But segmental duration is also associated with discourse-level structure: Relatively short unaccented expressions are more likely to be interpreted as references to entities previously mentioned in focused positions, whereas longer pitch-accented expressions tend to be interpreted as referring to previously unmentioned or non-focused entities (Dahan, Tanenhaus, and Chambers, 2002). Discourse context may therefore influence listeners' expectations about fine-grained acoustic-phonetic properties of an upcoming word, affecting how these cues are interpreted as a word unfolds.

A visual world study investigated the integration of prosodic and information-structural cues in spoken-word recognition. Participants' eye movements were monitored as they followed spoken instructions to move objects in a five-picture visual display. On each trial, the first instruction established a discourse context for the critical second instruction, which asked participants to manipulate the referent of a target word containing an onset-embedded word (e.g., *panda*). The initial syllable of the target word was manipulated to create two otherwise acoustically identical versions of the critical instruction: one short version, with flat F0 across the initial syllable, and one longer version, with an F0 rise of approximately 45 Hz. These two versions of the critical instruction were paired with two versions of the first instruction, which varied the thematic role of the target and competitor. The target was the theme in same-role (SR) contexts and the goal in different-role (DR) contexts.

The prosodic boundary hypothesis predicts increased competition from the onset-embedded competitor (e.g., *pan*) in long-target conditions, regardless of information structure. If, however, listeners interpret unaccented expressions as referring to entities previously mentioned as topics and accented expressions as signaling focus shift, SR contexts should reinforce the tendency for longer target words to elicit more competitor activation than shorter target words. Conversely, DR contexts should be associated with a reduced or even reversed effect: Although longer targets are consistent with pre-boundary lengthening, favoring the competitor, they are also consistent with a focus-shifted target word, disfavoring the competitor.

Consistent with the predictions of the information-structural account, the results showed a significant interaction between context type and target realization on competitor fixations. There were more fixations to the competitor in the SR-long condition than in the SR-short condition, and fewer fixations to the competitor in the DR-long condition than the DR-short condition. This interaction pattern suggests an effect of information

structure on the interpretation of the durational and pitch cues associated with long and short target words. The results provide evidence that the discourse context affects listeners' expectations about fine-grained acoustic-phonetic properties of the unfolding utterance, thus modulating the use of these cues in spoken word recognition.

Salverda, A.P., Dahan, D., and McQueen, J.M. (2003). The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension. *Cognition*, 90, 51–89.

Dahan, D., Tanenhaus, M.K., and Chambers, C.G. (2002). Accent and reference resolution in spoken language comprehension. *Journal of Memory and Language*, 47, 292–314.

Example item:

SR-context: Put the **panda** between the **pan** and the triangle...

DR-context: Put the **pan** between the **panda** and the triangle...

Target: Now put the **panda** above the circle.