CUE RELIABILITY AND ADAPTIVE RE-WEIGHTING IN SPOKEN WORD RECOGNITION

During spoken word recognition, listeners integrate bottom-up auditory cues from the signal with top-down cues (e.g., lexico-semantic context). Work on cue integration [e.g., 1, 2] has generally assumed that cue weights are static. In everyday language use, however, the relative reliability of cues (and thus their optimal weight [3,4]) can vary—e.g., between talkers. We explore whether listeners are sensitive to changes in the relative reliability of cues, and learn to re-weight cues accordingly. Specifically, we investigate the consequences of unexpectedly high degrees of cue conflict in recent input. We hypothesize that the cue that is more likely to have changed in reliability will be down-weighted over time [3, 4]. We investigate this for the integration of bottom-up phonetic and top-down semantic cues in spoken word recognition.
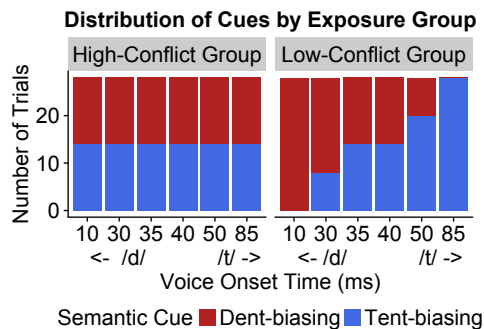

Figure 1: Between-Group design.

Methods. We presented listeners (N=106) with sentences like "When the ?ent in the [fender/campground] was noticed...", and they judge whether they heard "tent" or "dent" (following [5,6]). Two cues were varied: the phonetic continuum from /t/-/d/ (using VOT); and a binary semantic cue biases toward either "tent" or "dent". We divided subjects into two exposure groups. In the *High-Conflict group*, phonetic and semantic cues were uncorrelated, creating frequent conflict. In the *Low-Conflict group*, we decreased the number of conflicts (Fig 2).

Results. We replicated the main effects of both VOT ($\hat{\beta} = 0.1, p < 0.001$) and semantic cue ($\hat{\beta} = 0.9, p < 0.001$) found in previous work [5,6]. This confirms that both cues affect categorization. Critically, these effects interacted with the novel exposure manipulation: the semantic cue effect was larger in the Low-Conflict Group than in the High-Conflict Group (significant at all VOTs tested; $\hat{\beta}s >= 0.43, ps < 0.001$). This difference was driven by a three-way interaction between exposure group, trial, and semantic cue: semantic cues were down-weighted over time for the High-Conflict group, but not the Low-Conflict group ($\chi^2 = 12.34, p < 0.01$). Fig 2 illustrates this through fits from a generalized additive mixed model. A second study not reported here conceptually replicated this effect using a different design and stimuli. Interestingly, the effect of VOT on categorization changed little throughout the experiment.

Conclusions. Listeners re-weight cues depending on their reliability over time: listeners who encountered high levels of conflict between phonetic and semantic cues over time down-weight the semantic cues, relying instead primarily on phonetic cues. This effect emerged over time, suggesting that listeners cumulatively track the correlations of cues in their exposure to guide cue re-weighting. This highlights the adaptivity of spoken word recognition, and points to implicit learning processes that continuously update through life to support processing. It also raises questions about why listeners consider the phonetic cue (VOT) less likely to the be source of the cue conflict. One possibility is that listeners have stronger priors about the distribution of phonetic cues in experiments, compared to semantic cues.
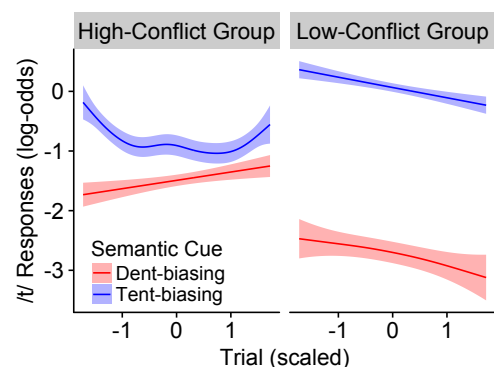

Figure 2: Weight of semantic cue (distance between lines) changes as function of group-specific exposure

[1] Oden & Massaro (1978) *Psych Review*. [2] Toscano & McMurray (2010) *Cognitive Science*. [3] Jacobs (2002) *TICS*. [4] Atkins et al. (2001) *Vision Research*. [5] Connine et al. (1991) *JML*. [6] OMITTED FOR REVIEW