Listeners Maintain Uncertainty About Acoustic Input

Wednesday Bushong (U of Rochester), Klinton Bicknell (Northwestern), T. Florian Jaeger (U of Rochester)
wbushong@ur.rochester.edu

Speech is a fast, temporally fleeting perceptual signal. In order to keep up with this constant stream of sensory input, listeners must compress the complex acoustic signal into more abstract representations, e.g. phonemes and words [1]. However, listeners may want to maintain uncertainty about such compressed representations, in case they need to update these beliefs at a later time. Consider the sentence "When the ?ent in the fender was well camouflaged, we sold the car." Depending on the voice onset time (VOT) of the ?-sound, listeners may interpret the word as either *tent* or *dent*. Later semantic context (e.g., *fender*), however, can disambiguate (e.g., towards *dent*). If listeners maintain uncertainty about linguistic input, their perception of *?ent* should be influenced by disambiguating context even when it comes after the critical word. Previous studies have found that listeners are in certain circumstances influenced by later context [2, 3, 4]. However, previous work has left open some important aspects of this maintenance process. **(Q1)** Are listeners capable of maintaining uncertainty for all linguistic input, or only ambiguous input? **(Q2)** Even if listeners are *capable* of rationally integrating evidence across all stimuli, are they more likely to make a perceptual *decision* before disambiguating context for less ambiguous stimuli?

Methods. We recruited participants from Amazon Mechanical Turk to perform a word labeling task (Expt 1: N=39, Expt 2: N=37). Participants listened to sentences like those in the table below, and indicate whether they heard the word *tent* or *dent*. We vary whether the word is disambiguated by context to either tent or dent and VOT (10, 40, 50, 60, 70, or 85ms). In Experiment 1, participants are required to listen to the entire sentence before they make a response. To address (Q2), in Experiment 2 participants can make a response anytime during the sentence (they can *choose* when to make a perceptual commitment).

Context	Sentence
Tent-biasing	When the [t/d]ent in the **forest** was well camouflaged, ...
Dent-biasing	When the [t/d]ent in the **fender** was well camouflaged, ...

Results. (Q1) If listeners rationally maintain uncertainty for all stimuli, we should observe context effects at all VOTs with similar magnitudes in log-odds space (see [4]). If this is an effect of *integration* rather than just *bias*, there should also be a main effect of VOT. We found a main effect of context, such that in contexts that disambiguated the word to *tent*, subjects were more likely to report that they heard *tent* (Expt 1, $\hat{\beta} = 0.22, p = 0.0019$; Expt 2 all trials, $\hat{\beta} = 0.81, p < 0.001$; Fig 1, 2). This effect was consistent across the VOT continuum: a simple effects analysis showed that the effect of context was significant at 40, 50, 70, and 85ms in Expt 1 ($\hat{\beta}$s $= 0.26$ to 0.51, ps < 0.05) and all VOTs in Expt 2 ($\hat{\beta}$s $= 0.31$ to 0.97, ps < 0.05). Critically, the context effect was *not* larger at more ambiguous VOTs but rather was relatively constant across the continuum (Fig 2). We did also find a main effect of VOT in both experiments (Expt 1: $\hat{\beta} = 0.17, p < 0.001$, Expt 2: $\hat{\beta} = 0.16, p < 0.001$; see Fig 1), suggesting that listeners are integrating context into their perceptual decision rather than using only context to make a response. **(Q2)** If listeners are less likely to wait for potential disambiguating information for less ambiguous stimuli, ambiguity should affect the likelihood of responding before the disambiguating word in Expt 2. We measured ambiguity as the distance of the VOT grand average from 50% t/d responses and found a significant effect on proportion of responses before the disambiguating word ($\hat{\beta} = 3.46, p < 0.001$; Fig 3).

Conclusions. Our results suggest that listeners have the ability to maintain uncertainty across all stimuli, not only ambiguous stimuli. However, although listeners integrate context when they hear it, they are less likely to decide to wait for potential disambiguating information when the perceptual signal is less ambiguous.

[1] Christiansen, M. & Chater, N. (2015) BBS. [2] Connine, C. et al. (1991) JML. [3] Szostak, C. M. & Pitt, M. A. (2013) AP&P. [4] Bicknell, K. et al. (2014) CUNY.

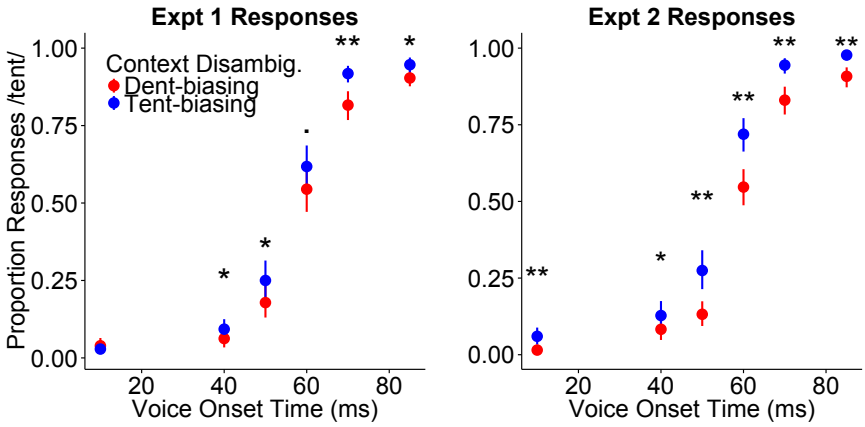

Figure 1: Expt 1 and 2 (all trials) responses by biasing context. Error bars are 95% confidence intervals after by-subject aggregation.

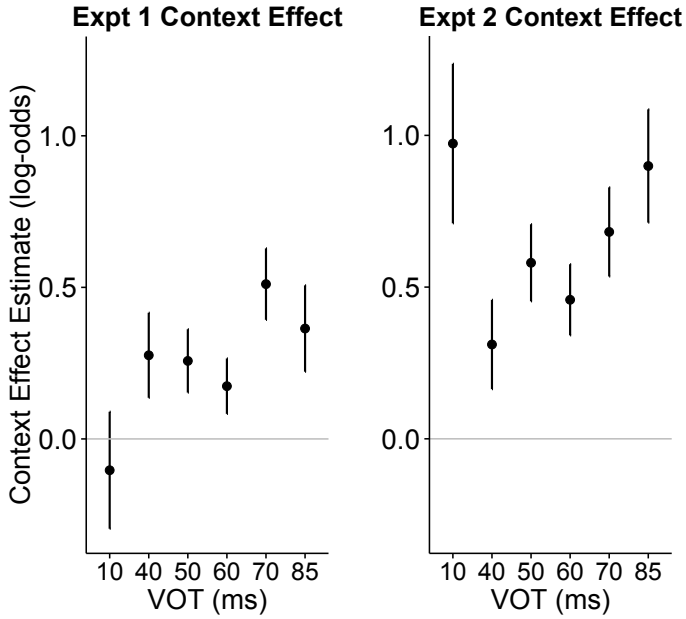

Figure 2: Expt 1 and 2 (all trials) context effects as estimated by our mixed logit model. Note that the most ambiguous VOT (60ms, see (a)) does not show a larger context effect than less ambiguous points.

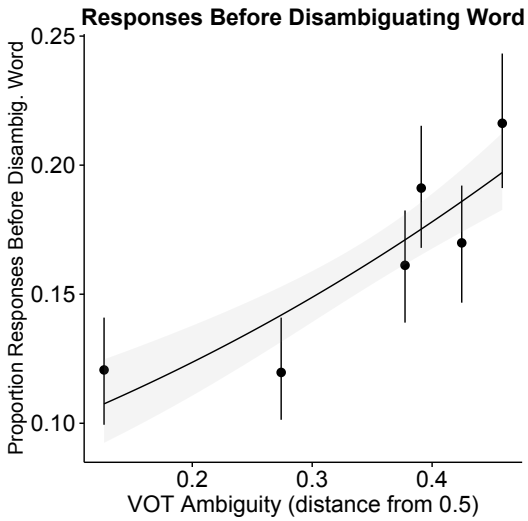

Figure 3: Experiment 2: likelihood of responding before disambiguating information by ambiguity of perceptual signal (points further to the right on the x-axis are less perceptually ambiguous). Error bars are 95% confidence intervals.