Listeners can maintain and rationally update uncertainty about prior words

Klinton Bicknell[1], Michael K. Tanenhaus[2,4], & T. Florian Jaeger[2,3,4]

[1] Department of Linguistics, Northwestern University

[2] Department of Brain & Cognitive Sciences, University of Rochester

[3] Department of Computer Science, University of Rochester

[4] Department of Linguistics, University of Rochester

Word count: 2970

Author Note

Please address all correspondence to Klinton Bicknell, at the Department of Linguistics, Northwestern University, 2016 Sheridan Road, Evanston, IL 60208-0854, USA.

E-mail: kbicknell@northwestern.edu

This research was supported by an NIH post-doctoral fellowship to KB (T32-DC000035 to the Center for Language Sciences at the University of Rochester), NSF CAREER IIS-1150028 to TJF, and NIH grant HD073890 to MKT. The authors thank Meredith Brown and Bozena Pajak for technical assistance in stimuli creation, and Dave Kleinschmidt for technical assistance in preparing the web-based experiment.

Abstract

Accurate word recognition is facilitated by context. Some relevant context, however, occurs after the word. Efficient use of such "right context" requires comprehenders to *maintain uncertainty* about the word, still allowing for consideration of multiple possible alternatives when they encounter relevant right context. However, influential models suggest that uncertainty is not maintained in this way. A classic study (Connine et al., 1991, Experiment 1) examined right context effects using word pairs that differed in voicing by manipulating VOT (e.g., dent/tent). The results were interpreted as evidence for limited uncertainty maintenance. Right context effects were limited to fewer than six syllables downstream and even then were only found for highly ambiguous VOTs near the category boundary. With small modifications in procedure and analysis, we report that uncertainty is maintained for at least six to eight syllables and equally so for the entire VOT continuum (rather than only ambiguous cases). We show that an ideal recognizer, which optimally combines acoustic information with right context, correctly predicts our results. This suggests that, at least under some conditions, listeners combine acoustic information with right context rationally.

Keywords: rational comprehension; maintaining uncertainty; VOT; word identification; right context

Listeners can maintain and rationally update uncertainty about prior words

1. Introduction

Language comprehension requires inferring linguistic structure from perceptual input. Following Marslen-Wilson's seminal work (Marslen-Wilson, 1973, 1975), models of language comprehension have adopted some variant of what Just and Carpenter (1980) labeled the "immediacy assumption": input is fully processed, i.e., integrated into representations at multiple levels, immediately. The immediacy assumption is typically accompanied by a second assumption: once input is integrated, uncertainty about how to categorize the input is rapidly resolved. Motivation for immediate uncertainty reduction has roots in classic accounts of short-term memory in which sensory memory rapidly decays (e.g., Sperling, 1960).

In speech perception, this assumption, though often implicit, has been enormously influential (for review, see Christiansen & Chater, accepted for publication). It is reflected in standard views of categorical perception: whereas listeners are initially sensitive to within-category differences in acoustic cues such as VOT, sensitivity rapidly decays as inputs are parsed into categorical representations (e.g., Goldstone & Hendrickson, 2010). Well-known models of spoken word recognition also make this assumption. In the original cohort model (Marslen-Wilson & Welsh, 1978), lexical candidates are completely eliminated once they become inconsistent with the input. The influential TRACE model (McClelland & Elman, 1986), which embodies principles of interactive activation, implements a somewhat weaker version of immediate uncertainty reduction with lateral inhibition at the phoneme and word levels.

An alternative view is that the systems underlying comprehension are: (1) less directly constrained by resource limitations; and (2) organized to combine information in the perceptual signal with information in the linguistic context in ways that are probabilistically justified (e.g., Clayards et al., 2008; Feldman et al., 2009; Levy et al., 2009; Kleinschmidt & Jaeger, 2015; Norris & McQueen, 2008). For example, a comprehension system that makes rational use of bounded short-term memory should not immediately discard all uncertainty. Instead, uncertainty should be maintained for as long as the *expected* advantages (e.g., more accurate word recognition) outweigh the disadvantages (e.g., attentional and other resource costs and interference from incoming input; for related ideas in reading, see Bicknell & Levy, 2010; Lewis et al., 2013). Moreover, uncertainty maintenance is required to efficiently deal with inter-talker variability (Kleinschmidt & Jaeger, 2015, p180-181).

Evidence that uncertainty about speech input is not immediately discarded comes from effects of “right context”, in which comprehenders use subsequent context to inform word and segment recognition (Grosjean, 1985; Bard et al., 1988; Connine et al., 1991; see Dahan, 2010, for a review). For example, using pairs of words that differed in initial-consonant voicing, but shared the next few syllables, e.g., “parakeet “and “barricade”, McMurray, Tanenhaus & Aslin (2009) showed that gradient effects of VOT are maintained and used when listeners encounter information several syllables downstream. However, many models relax the immediate uncertainty reduction assumption to allow uncertainty maintenance within a word..

Despite the importance of the assumption of immediate uncertainty reduction in guiding psycholinguistic theory, we know of only two published studies that directly investigated how long uncertainty is maintained beyond the word (Connine et al., 1991; Szostak & Pitt, 2013). In a pioneering study, Connine et al. (1991, Experiment 1) found that uncertainty maintenance seems strictly limited in two ways: (1) It is short-lived, persisting for less than six syllables after word offset; and (2) It is restricted to highly ambiguous VOTs close to the category boundary. While these conclusions are incompatible with the immediate uncertainty reduction assumption, they are also incompatible with rational accounts. Instead, they are consistent with a slightly weakened version of the immediate uncertainty reduction assumption: whereas left context effects are pervasive and part and parcel of "normal language processing", right context effects are limited to special cases approaching maximal ambiguity and are very limited in time. Here, we present new data and analyses that cast doubt on limitations (1) and (2) and instead support a rational account. We begin with two concerns about Connine et al.'s design and analysis that compromise their conclusions.

2. Connine et al. (1991)

Connine et al. investigated the perception of words with voiceless or voiced plosives onsets ([t]ent–[d]ent). The target stimulus was embedded in a sentence frame in which subsequent context, 3 (near) or 6–8 (far) syllables after word offset, supported the voiced or voiceless interpretation (Figure 1a). The plosive's VOT varied from values prototypical for voiced [d] to prototypical for voiceless [t]. The results (Figure 1b) suggested that right context influenced listeners' identification of the target when it occurred 3 syllables later, but not 6–8. Connine et al. concluded that uncertainty is resolved within 6–8 syllables. Moreover, they suggested that near-condition context effects were restricted to VOTs with maximal uncertainty, i.e., near the category boundary (also see, Szostak & Pitt, 2013, p. 1539).

(a)	[dent-biasing, near]	When the ___ in the fender was well camouflaged, we sold the car.
	[tent-biasing, near]	When the ___ in the forest was well camouflaged, we began our hike.
	[dent-biasing, far]	When the ___ was noticed in the fender, we sold the car.
	[tent-biasing, far]	When the ___ was noticed in the forest, we stopped to rest.

(b)

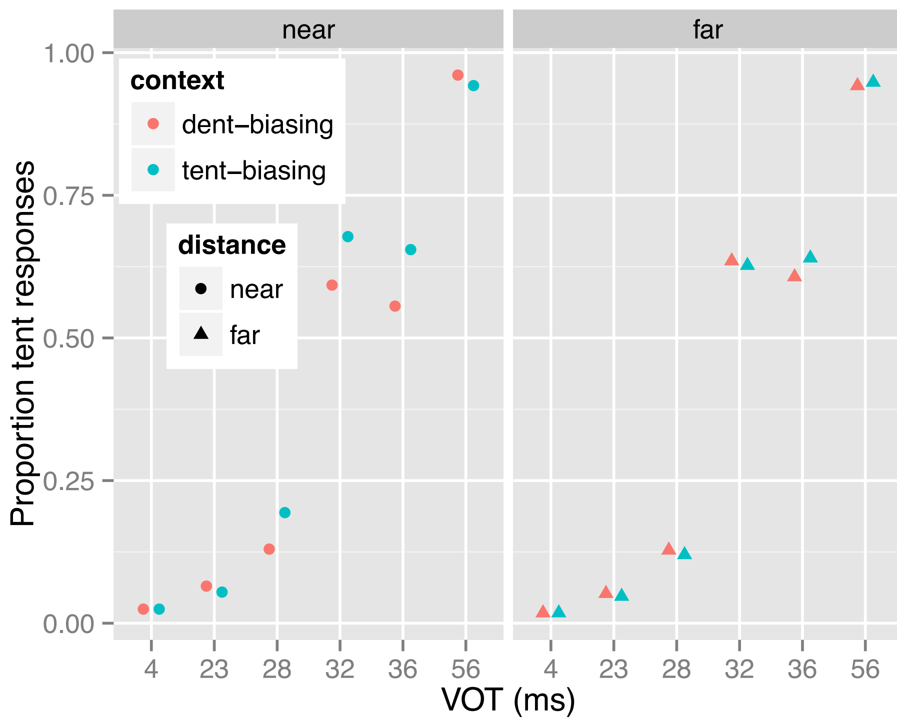

Figure 1. Experiment 1 from Connine et al. (1991). (a) Four experimental conditions for sentence frames, presented with a sample item. Six steps on a VOT continuum from 4 ms (voiced, [dent]) to 56 ms (voiceless, [tent]) were inserted into each sentence frame condition. (b) Mean proportions of ‘tent’ responses in each combination of VOT and sentence frame (replotted). A reliable difference between dent-biasing and tent-biasing subsequent contexts was found only for the *near* condition, and this difference was driven by VOTs near the category boundary (28, 32, and 36 ms). Note the discontinuous x-axis.

A closer look at the Connine et al. procedure and analysis suggests the results might actually be compatible with a rational ideal recognizer account. The absence of context effects in the far condition could be an artifact of the procedure, which allowed participants to respond during the sentence. In fact, 84% of responses preceded the biasing information in the far condition (versus 15% in the near condition[1]). Therefore the study might not have had sufficient power to detect right context effects at long lag. Indeed, a recent study on fricative perception that avoided this problem (Szostak & Pitt, 2013) found uncertainty maintenance more than 3 syllables downstream. Rather than attributing this difference to the massive data loss in the far condition in Connine et al (1991), Szostak and Pitt (2013) hypothesized that "memory trace of fricatives may be longer lasting than the memory trace for plosives". Our first goal is to test whether VOT uncertainty is maintained 6–8 syllables after word offset by requiring participants to wait until sentence offset to make an identification response.

Second, Connine et al.'s conclusion that context effects are limited to VOTs in the most ambiguous range is based on a statistical analysis showing that context effects are larger in proportion space for VOTs near the category boundary. However, we show that an ideal recognition system, which combines all acoustic information in the signal with all information in the subsequent context, also predicts that the effects are largest near the category boundary, even though right context is used for *all* tokens. Our second goal is to derive and test quantitative predictions to determine (a) whether listeners maintain and update uncertainty for all tokens or only for maximally ambiguous ones and (b) if uncertainty is maintained for all tokens and, information is combined optimally.

---

[1] This difference may be due to how many syllables downstream listeners expect critical information about a word. Alternatively, participants might simply have responded as fast as possible after the critical word (giving them access to the critical subsequent context in the near but not the far condition). We return to these possibilities

3. Experiment

To achieve these goals we designed a study with materials and a task modeled closely on Connine et al., Experiment 1. To determine whether listeners can still maintain and update uncertainty 6–8 syllables after word offset, we modified the procedure so that participants could not respond until the end of the sentence, so that we have sufficient power to detect effects of right context with 6-8 syllable delays. To achieve the second goal we derive quantitative predictions for an ideal recognizer and compare them with predictions of a model in which right-context effects are restricted to the most ambiguous tokens, using statistical analyses examining the effects of VOT and right context in log-odds space.

3.1. Method

3.1.1. Participants. Forty-eight workers on Amazon Mechanical Turk participated in the experiment. Two participants were excluded for performing the experiment twice.

3.1.2. Materials. We used materials modeled after Connine et al. (1991, Experiment 1). We used a tent–dent continuum and seven sets of sentence frames taken from Connine et al. (Appendix A, sets 1–6 and 8). Each set had four conditions identical to those in Connine et al. (Figure 1a), yielding a total of 28 sentence frames. For each set, the material preceding the target word was identical.

A female speaker in a noise-attenuated booth recorded the sentence frames with the target word biased by the right context. To create identical pre-target frames across the four versions, we concatenated one version of the first word for each set (‘when’ in example 1) with a single recording of ‘the’, used for all sentences.

---

in the discussion. In any case, this paradigm severely limits the power to detect whether listeners maintain uncertainty.

Following Connine et al., we created the tent–dent VOT continuum by selecting a recording of *dent* with a relatively long vowel, and then successively replacing more of the waveform immediately following the release burst with the waveform that immediately followed the burst in *tent* (Figure 2). Based on a norming study in which participants identified members of this continuum without context, we selected six VOT values: two unambiguous endpoints (10 and 85 ms) and four values around the category boundary (40, 45, 50, 55 ms). The 28 target frames were combined with the 6 VOTs to create 168 sentences. All recordings are available at https://www.hlp.rochester.edu/resources/BicknellTanenhausJaeger/.

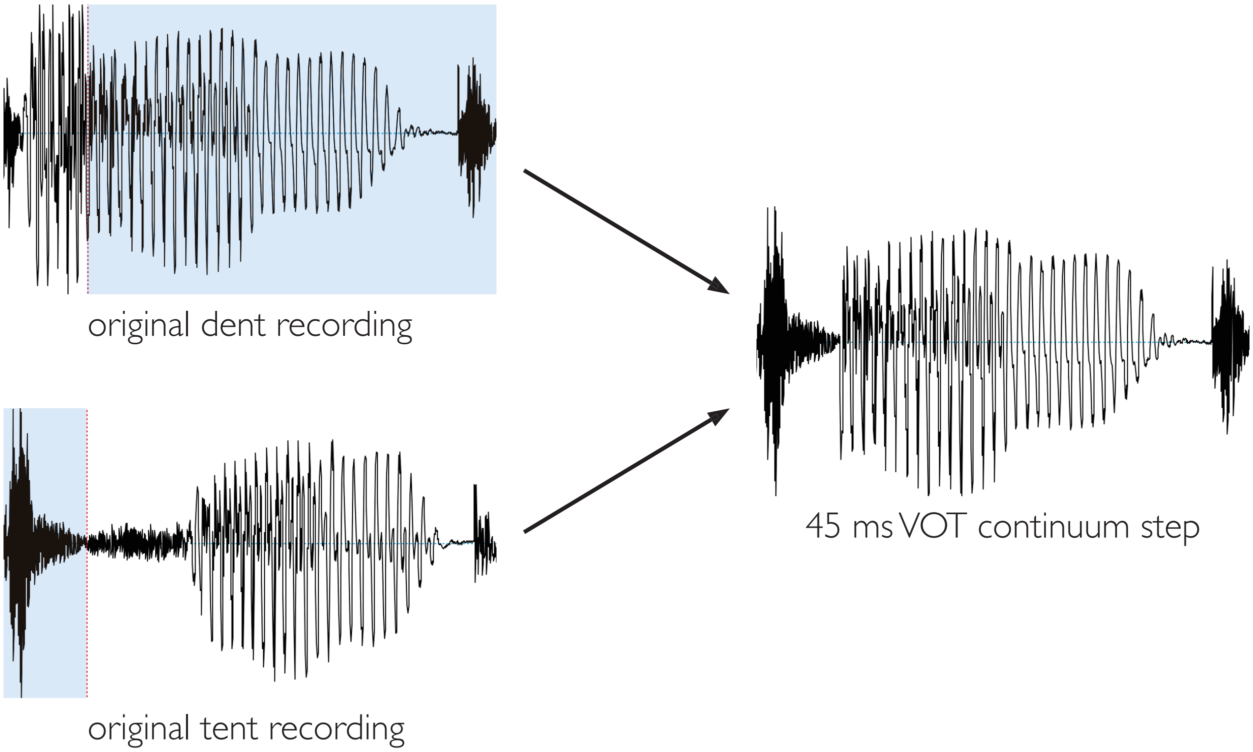

Figure 2. Illustration of VOT continuum step creation. From two original recordings of dent and tent (left), each starting with the burst from the initial [t]/[d], we created a VOT continuum step waveform with a particular VOT (here 45 ms, on right) by concatenating the first 45 ms of the tent waveform (in blue, lower left) with the remainder of the dent waveform starting at 45 ms after the burst (in blue, upper left).

3.1.3. Procedure. After giving informed consent, participants listened to each of the sentences (following Connine et al.), in individually randomized order. After each sentence, they judged whether it contained *dent* or *tent*.

3.1.4. Analysis. We excluded five participants whose responses showed no effect of VOT (at the $p<0.1$ level), suggesting that they did not understand the task. We analyzed the remaining response data using logistic mixed-effects regression (for an introduction, see Jaeger, 2008) via the *lme4* package (Bates et al., 2015). Fixed effects included the three design variables – *bias*, *distance*, and *VOT* – and their interactions. Bias and distance were deviation-coded, and VOT was continuous. For reasons described below, we also included an orthogonal quadratic effect of VOT and its analogous interactions. We included random intercepts and random slopes for all fixed effects by participants and items. To achieve convergence, we omitted random correlations except between VOT and the intercept. We assess effects via the likelihood ratio test, comparing the full model to models without each fixed effect.

3.2. Results

The response data are visualized in Figure 3.

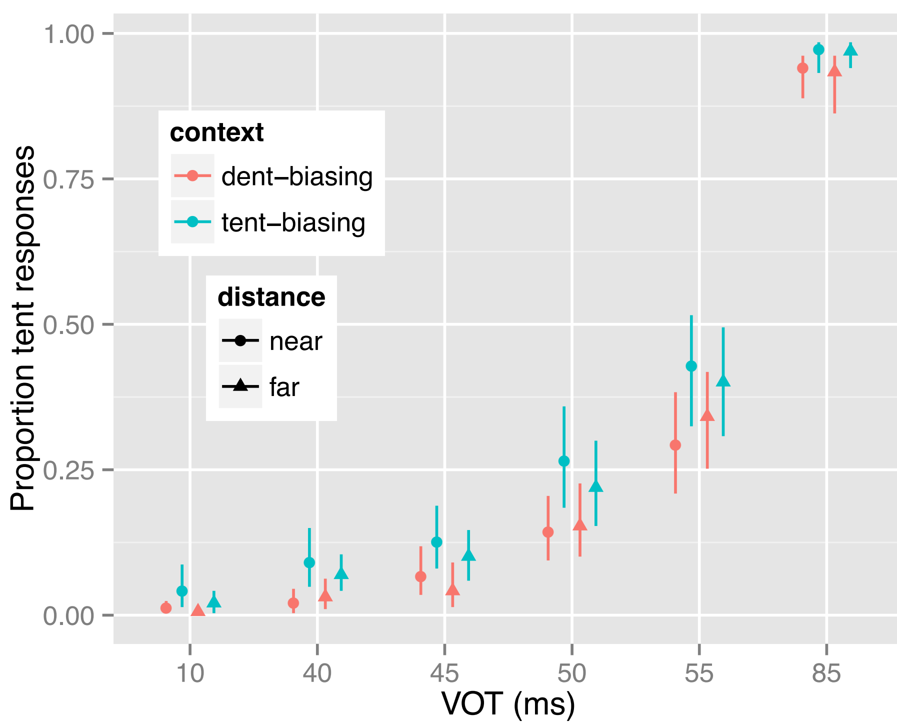

Figure 3. Proportion of ‘tent’ responses in each condition. Error bars show 95% confidence intervals, BCa bootstrapped over participant means. Note the discontinuous x-axis.

As expected, VOT had the largest effect, ($p<0.001$): with increasing VOT listeners were more likely to respond *tent* than *dent.* Replicating Connine et al., we find a main effect of subsequent contextual bias ($p<0.001$): listeners were more likely to respond *tent* for tent-biasing contexts than dent-biasing contexts. However, whereas Connine et al. found bias effects only for the near condition, we find reliable effects of contextual bias for both near ($p< 0.001$) and far (6–8 syllable) conditions ($p<0.01$), with no reliable interaction between distance and bias ($p=0.5$). After 6–8 syllables, listeners can still use linguistic context to update uncertain beliefs about a prior word.

3.2.1. Testing quantitative predictions of an ideal recognizer. We contrast an ideal recognizer with a model in which effects are restricted to tokens near the category boundary, which we will term the ambiguity model. The ideal recognizer predicts that acoustic evidence (here VOT) and subsequent context should have constant (additive) effects in log-odds space (See the derivation in the Appendix). A constant-sized effect in log-odds will be largest around 0.5 in probability space (see Jaeger, 2008). Thus, the ideal recognizer, like the ambiguity model, predicts that the effect of subsequent context will be largest in probability space for tokens near the category boundary. This was the case for Connine et al.'s data, and our data as shown in Figure 3. However, the two accounts make different predictions for VOTs further from the category boundary. The ambiguity account predicts context effects only for VOTs near the category boundary. In contrast, the ideal recognition account predicts that whereas the size of the context effect will grow smaller in probability space, it will remain constant in log-odds space. As Figure 4 illustrates, we find no evidence for a diminished effect of context in log-odds space for the continuum endpoints (if anything the data numerically point in the opposite direction).

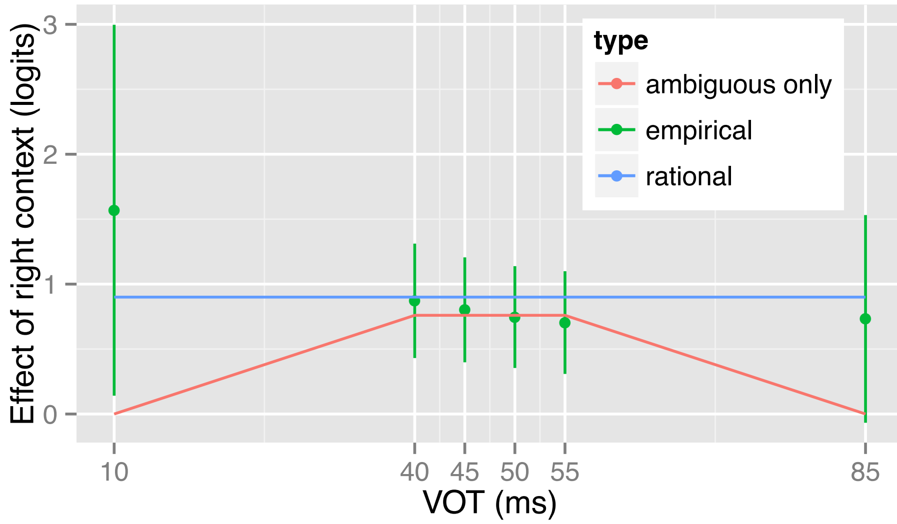

Figure 4. Effect of subsequent contextual bias as estimated by our mixed-effects regression model. Shown are the present data with error bars giving 95% CIs on effect size (green); a schematic prediction of fully rational integration, which predicts a constant effect across all VOTs (blue); a schematic prediction of the ambiguity model, which predicts no effect of bias for unambiguous stimuli (red).

These predictions map straightforwardly onto (mixed-effects) logistic regression, which operates in log-odds space. An account in which effects of subsequent context are strongest for maximally ambiguous VOTs predicts an interaction of subsequent context with quadratic VOT, such that contextual effects are smaller (or even nonexistent) for VOTs further from the category boundary. Ideal recognition models predict no such interaction, because the effect of context should be constant in log-odds for all VOTs.

We therefore included a quadratic VOT term and all its interactions in the model. There were no reliable interactions with quadratic VOT, including the crucial interaction with bias

(p=0.4). Therefore the results are compatible with right context having a constant effect size. This result is fully consistent with an ideal recognizer in which listeners maintain and update uncertainty for all phonetic tokens to approximate optimal combination of acoustic cues with subsequent context. It is, however, inconsistent with the ambiguity model.

4. Discussion

We found no evidence that maintenance of uncertainty for VOT is limited to fewer than six syllables or is restricted to maximally ambiguous tokens. Rather, at least under the conditions we tested, listeners can maintain and update uncertainty about a word for at least 6–8 syllables after its offset, the longest lag tested. Further, our analyses suggest that listeners can maintain uncertainty about all words, regardless of their level of ambiguity, and rationally combine the information present in the acoustics with the information in the right context.

Therefore our results support a model in which listeners combine acoustic information with linguistic context rationally, even if the relevant linguistic context occurs substantially later. Therefore listeners do not always immediately discard the information present in the acoustic data but rather maintain access to this information across subsequent words.

While the current results show that listeners *can* rationally maintain and update uncertainty about previous words, the extent to which they generally do so, remains an open question. The paradigm used here (and by Connine et al., 1991; Szostak & Pitt, 2013) is atypical in that listeners judge the same word many times and also know in advance that there are two alternatives. It is thus an open question whether the rational integration we see here is a typical property of comprehension. It will be important, then, to use more naturalistic settings to examine: (a) how far downstream relevant linguistic context is usually observed; and (b) how long listeners typically maintain uncertainty; and (c) whether these two are related.

Another important question to address is how precisely listeners maintain uncertainty: do they maintain subphonemic information or merely track their degree of uncertainty about the phonemic category? The latter may tax sensory memory less. There is, however, reason to believe that listeners maintain subphonemic information: listeners can store subphonemic information about specific talkers and talker groups over extended periods (e.g., Walker & Hay, 2011; Eisner & McQueen, 2006; Goldinger 1996; Johnson et al., 1999; reviewed in Weatherholtz & Jaeger, submitted). Maintaining this information allows listeners to deal with inter-talker variability in the realization of phonetic units (for discussion, see Kleinschmidt & Jaeger, 2015). It also facilitates talker recognition and inferences about talker's social identity (ibid).

The existence of such knowledge alone would be puzzling if language comprehension was so severely constrained by, for instance, memory limitations (cf. Christiansen & Chater, accepted for publication) that all uncertainty about linguistic categories had to be immediately discarded. Instead, the current results, along with research on talker-specific expectations, suggest the language processing system is less constrained than previously thought, and uses its available resources to achieve performance that approximates that of an ideal recognizer.

Appendix

This appendix derives the prediction of an ideal recognizer that perceptual evidence and context should have additive effects in log-odds. The probability of a word *w* being *tent* given just a particular subsequent linguistic context *c* is given by Bayes rule:

(1) $$p(w = \text{tent}|c) = \frac{p(c|w = \text{tent})p(w = \text{tent})}{p(c)}$$

Conditioning all terms on the perceptual evidence *e* for the word *w* yields the probability that an ideal observer should believe w is *tent*, given both acoustics and right context.

(2) $$p(w = \text{tent}|c, e) = \frac{p(c|w = \text{tent}, e)p(w = \text{tent}|e)}{p(c|e)}$$

Assuming the context *c* is conditionally independent from the evidence *e* for word *w* given that we know word *w* (i.e., speakers do not change what subsequent context they produce based on the perceptual realization *e* of *w*), we can simplify:

(3) $$p(w = \text{tent}|c, e) = \frac{p(c|w = \text{tent})p(w = \text{tent}|e)}{p(c|e)}$$

Finally, to determine the log-odds that *w* is *tent* (as opposed to *dent*) for an ideal observer, we first find the odds by taking the ratio of the two posterior probabilities

(4) $$\frac{p(w = \text{tent}|c, e)}{p(w = \text{dent}|c, e)} = \frac{\frac{p(c|w = \text{tent})p(w = \text{tent}|e)}{p(c|e)}}{\frac{p(c|w = \text{dent})p(w = \text{dent}|e)}{p(c|e)}}$$

(5) $$= \frac{p(c|w = \text{tent})}{p(c|w = \text{dent})} \frac{p(w = \text{tent}|e)}{p(w = \text{dent}|e)}$$

then taking the logarithm of both sides:

(6) $$\underbrace{\log \frac{p(w = \text{tent}|c, e)}{p(w = \text{dent}|c, e)}}_{\text{log-odds of tent}} = \underbrace{\log \frac{p(c|w = \text{tent})}{p(c|w = \text{dent})}}_{\text{constant given context}} + \underbrace{\log \frac{p(w = \text{tent}|e)}{p(w = \text{dent}|e)}}_{\text{constant given evidence}}$$

Thus, the log-odds of *tent* given both perceptual evidence and subsequent context equals the sum of one term depending on subsequent context and one depending on perceptual evidence (and mutatis mutandis for *dent*). For an ideal recognizer, context and acoustics have additive effects on log-odds.

References

Bard, E. G., Shillcock, R. C., & Altmann, G. T. (1988). The recognition of words after their acoustic offsets in spontaneous speech: Effects of subsequent context. *Perception & Psychophysics*, *44*(5), 395-408.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015.) lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. http://CRAN.R-project.org/package=lme4

Bicknell, K., & Levy, R. (2010). A rational model of eye movement control in reading. In *Proceedings of the 48th annual meeting of the Association for Computational Linguistics* (pp. 1168-1178). Association for Computational Linguistics.

Christiansen, M. H. & Chater, N. (In press.) The Now-or-Never Bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*(3), 804-809.

Connine, C. M., Blasko, D. G., & Hall, M. (1991). Effects of subsequent sentence context in auditory word recognition: Temporal and linguistic constrainst. *Journal of Memory and Language*, *30*(2), 234-250.

Dahan, D. (2010). The time course of interpretation in speech comprehension.*Current Directions in Psychological Science*, *19*(2), 121-126.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *The Journal of the Acoustical Society of America*, *119*(4), 1950-1953.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review*, *116*(4), 752–782.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, memory, and cognition*, *22*(5), 1166–1183.

Goldstone, R. L., & Hendrickson, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69-78.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception & Psychophysics*, *38*(4), 299-310.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434-446.

Johnson, K., Strand, E. a, & D’Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359–384.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*(50), 21086-21090.

Lewis, R. L., Shvartsman, M., & Singh, S. (2013). The adaptive nature of eye movements in linguistic tasks: How payoff and architecture shape speed-accuracy trade-offs. *Topics in Cognitive Science*, *5*(3), 581-610.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244*, 522–523.

Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. *Science, 189*, 226–228.

Marslen-Wilson, W. D., & Welsh, A. (1978.) Processing interations and lexical access during word recognition in continuous speech. *Cognitive Psychology 10*, 29–63.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1-86.

McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2009). Within-category VOT affects recovery from “lexical” garden-paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language*, *60*(1), 65-91.

Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29.

Szostak, C. M., & Pitt, M. A. (2013). The prolonged influence of subsequent context on spoken word recognition. *Attention, Perception, & Psychophysics*,*75*(7), 1533-1546.

Walker, A., & Hay, J. (2011). Congruence between “word age” and “voice age” facilitates lexical access. *Laboratory Phonology*, *2*(1), 219–237.

Weatherholtz, K. and Jaeger, T. F. (submitted). Speech perception and generalization across talkers and accents. For *Oxford Research Encyclopedia of Linguistics*.