

N400 Evidence for Parallel Lexical Predictions

There is a growing consensus, stemming from both behavioral and neurolinguistic data, that comprehenders generate probabilistic predictions about upcoming words (Altmann & Kamide, 1999; Kuperberg & Jaeger, 2015). For example, DeLong, Urbach, & Kutas (2005) demonstrated that N400 amplitudes vary gradiently with the cloze probability of a target noun. They take this as evidence that word representations are pre-activated in proportion to their probability of being seen next (given the context). Computational theories of prediction have often assumed that this entails parallel lexical pre-activation with multiple lexical items being predicted in parallel, each with different strengths (or probabilities) (Levy, 2008; Smith & Levy, 2013). However, Huettig & Mani (2015) have argued that much of the neurological evidence for prediction relies upon unnatural experimental contexts in which comprehenders are disproportionately led to expect (and then indeed receive) a highly probable word, i.e. that prediction may not be as prevalent or as gradient as is often suggested, but rather occurs only in specialized circumstances.

Here we used ERPs to explicitly ask whether multiple probable candidate words are predicted in parallel in medium-constraint contexts. This account is widely assumed, but has not been explicitly tested in previous ERP paradigms of gradient lexical prediction. This is because, with the exception of zero cloze probability words (see Federmeier et al., 2007), previous ERP studies have effectively confounded the probability of a given critical word (its cloze probability) with the probability of the most expected word, given its preceding context (its lexical constraint). For example, DeLong, Urbach, & Kutas' study might include a 40% cloze target which is the most probable word given the context, but not a 40% cloze target in a context where the remaining 60% probability is given to a more-probable word.

We carried out an ERP study in which participants ($n=32$) read short story contexts that generated predictions for at least two probable words. These contexts were then continued with either the Best (i.e. most probable) or the Second Best (i.e. second most probable) word (Table 1). To further study the effect of a wide range of constraint values on prediction, contexts spanned constraint values from 36-92%. The same target words were also presented in Unconstraining contexts which generated no strong predictions (all constraints $\leq 36\%$). Under a parallel gradient account of prediction, we expect to see facilitation of the Second Best words in proportion to their probability, even when contextual constraint is kept constant, reflected by a three-way dissociation of N400 amplitudes: Best < Second Best < Unpredicted. Under a non-parallel account of prediction, and/or if previous findings were due to the confound between cloze probability and contextual constraint, only the most predictable word should receive facilitation, predicting only a two-way dissociation of N400 amplitudes: Best < Second Best = Unpredicted.

Using cluster-based permutation tests to control for multiple comparisons, we find a three-way dissociation of N400 amplitudes: Best < Second Best < Unpredicted (match Best) = Unpredicted (match Second Best) (Table 2, Figure 1). This finding supports a parallel gradient account of prediction in which multiple words can be pre-activated simultaneously in proportion to their probability of occurrence in context.

In future work, we plan to use single-trial analysis to investigate in more detail whether N400 amplitudes in our experiment varied with the cloze probability of the specific target word and, furthermore, whether there is any discernable later "penalty" for not being the best completion (e.g. does a 40% cloze Second Best target behave identically to a 40% cloze Best target?). This data will thus further allow us to test whether the potential for competition between lexical alternatives influences semantic processing, as indexed either by the N400 or potentially by other ERP components (Wlotko & Federmeier, 2012).

Table 1.

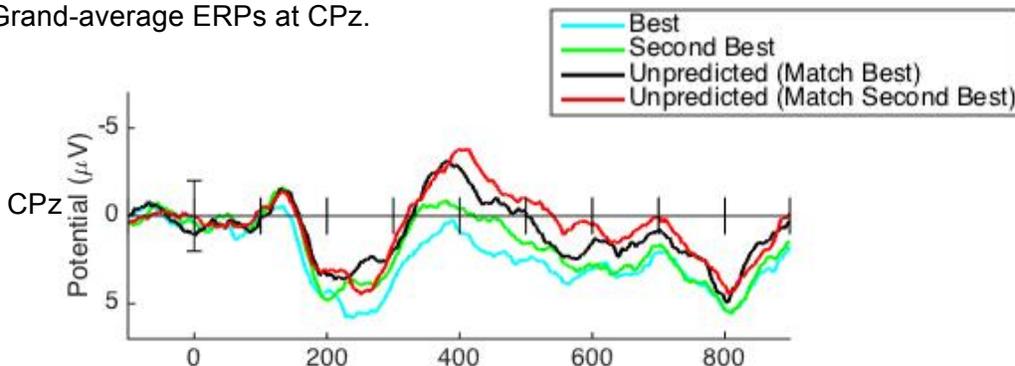
	Context	Target	Avg. Cloze	Cond. Label
Constraining	Stephen wanted to do something special for his girlfriend. He decided to make her a hand-made card. On it, he drew some...	hearts	57%	Best
		flowers	16%	Second Best
Unconstraining	Malcom always doodled in class. He took out a fresh sheet of paper. On it, he drew some...	hearts	4%	Unpredicted (match Best)
		flowers	3%	Unpredicted (match Second Best)

Table 2.

Comparison	Best vs. SecondBest	SecondBest vs. Unpredicted (Best)	SecondBest vs. Unpredicted (SecondBest)
p value	.01	.02	0.0004

Figure 1.

Grand-average ERPs at CPz.



References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8), 1117–1121.

Federmeier, K. D., Wlotko, E. W., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84.

Huetting, F., & Mani, N. (2015). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, 31(1), 19–31.

Kuperberg, G. R., & Jaeger, T. F. (2015). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32–59.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177.

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, 128(3), 302–319.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, 62(1), 356–366.