**Prediction in the Maze: Probabilistic pre-activation and the English *a/an* contrast**

E. Matthew Husband (University of Oxford)
matthew.husband@ling-phil.ox.ac.uk

The past two decades has seen a resurgence of interest in the idea that comprehenders routinely engage in prediction of upcoming linguistic content. Experimental studies exploiting morphosyntactic and phonotactic constraints on a word form preceding a high cloze target word have been key to underpinning predictive accounts of comprehension by providing unambiguous evidence for predictive processes. Most of this research uses ERPs, whose high temporal resolution and word-by-word presentation have been critical, especially when investigating phontactic/orthographic constraints like the English *a/an* contrast which is highly local, occurring between two adjacent words. Investigating these tight sequential contrasts with other traditional methods is difficult. *Self-paced reading* is unable to isolate focal effects to a specific word due to spillover effects (Mitchell, 1984), and, in *eye movements*, short functional words are often skipped and upcoming information may be available in parafoveal preview even when such words are fixated, requiring careful experimental manipulation (Rayner, 2009).

Given its much more focal response (Witzel, Witzel, & Forster, 2012; Witzel & Forster 2014), the *maze task* may address some of the potential shortcomings of these other methods and provide a cheap and easy alternative methodology to study early cues to prediction. In the maze task, each sentence is presented to participants as a sequence of choices between two words: one word is the correct continuation of the sentence, the other is a distractor. Automated distractor selection has made the maze task easier to implement (Boyce, Futrell, & Levy 2020), but it remains to be shown that this task can be sensitive to these early predictive cues.

**Experiment**. To address this, 80 sentence contexts with two possible continuations, an expected and an unexpected English *a/an* + noun combination from Nieuwland, et al.'s (2018) replication of DeLong, Urbach, & Kutas (2005) were counterbalanced and presented to 40 native English speakers from Prolific using IbexFarm. Distractor words were automatically generated using A-maze (Boyce, Futrell, & Levy 2020) with the Gulordava language model (Gulordava, et al., 2018). Participants were allowed to continue if they selected the wrong continuation, and a yes/no comprehension question followed one quarter of the sentences.

**Results**. Comprehension accuracy was high (avg. 91%) and error rates per word were low (avg. 2.8%; target articles: 3.3%, target nouns: 2.0%; see Table 1). Error responses were removed prior to reading time analysis. Reading times are shown in Figures 1 and 2A. LME models revealed that the unexpected condition was significantly slower than the expected condition on both the noun (294 msec, $t = 11.4$, $p < .001$) and CW+1 (42.6 msec; $t = 3.579$, $p < .001$) but also on the article preceding the noun (28 msec, $t = 3.421$, $p = .001$). Similar results were obtained using either noun cloze (article: $t = -2.885$, $p = .006$; noun: $t = -13.36$, $p < .001$; CW+1: $t = -3.555$, $p < .001$) or article cloze (article: $t = -3.937$, $p < .001$; noun: $t = -11.63$, $p < .001$; CW+1: $t = -3.635$, $p < .001$) as a continuous predictor. No significant differences were found in the regions prior to the article.

Figure 2A suggested that slower readers showed more of a predictive effect. Average reading time was calculated for each participant based on the average reading times of the three words prior to the article (CW-3, CW-2, CW-1) and added as a predictor to the models above. Significant interactions between participant average reading times and expectation were found on both the article (by expectation: $t = 3.52$, $p < .001$; by noun cloze: $t = -3.24$, $p = .002$; by article cloze: $t = -4.45$, $p < .001$) and the noun (by expectation, $t = 2.14$, $p = .040$; by noun cloze: $t = -2.34$, $p = .025$; by article cloze: $t = -1.84$, $p = .075$), shown in Figure 2B.

**Discussion**. By providing a focal reading time measure, the maze task was able to reveal effects of expectation on the *a/an* contrast prior to the noun and may be a useful alternative methodology for investigating predictive effects in comprehension. Interesting, early predictive effects of article form were greater for slower readers, suggesting that probabilistic pre-activation of the expected word to the level of phonological form may take time to emerge. It may be that the fragility of prior results using *a/an* are driven at least in part by temporal processing factors, consistent with prediction-as-production theories (Ito, et al, 2020).

**Example stimuli** [avg. article cloze: $73.9_{exp}/7.8_{unexp}$; avg. noun cloze: $81.8_{exp}/10.1_{unexp}$]

1. The old wives' tale says that if you want to keep the doctor away then you should eat <u>an apple$_{expected}$/a carrot$_{unexpected}$</u> a day. (art. cloze 82/0, noun cloze: 97/0)
2. For the snowman's eyes the children used two pieces of coal, and for its nose they used <u>a carrot$_{expected}$/an apple$_{unexpected}$</u> from the fridge. (art. cloze 98/0, noun cloze: 100/0)

**Table 1**: Average error rates by word. Standard errors by subject are given in paratheses.

| | CW-3 | CW-2 | CW-1 | art | n | CW+1 | CW+2 | CW+3 |
|---|---|---|---|---|---|---|---|---|
| **Expected** | 2.7 (2.7) | 3.1 (2.9) | 3.1 (2.9) | 4.0 (3.3) | 1.6 (2.6) | 3.0 (2.8) | 3.0 (2.9) | 2.3 (2.5) |
| **Unexpected** | 3.1 (2.9) | 3.1 (2.9) | 2.4 (2.6) | 2.7 (2.7) | 2.5 (2.6) | 2.4 (2.6) | 3.9 (3.2) | 2.0 (2.3) |

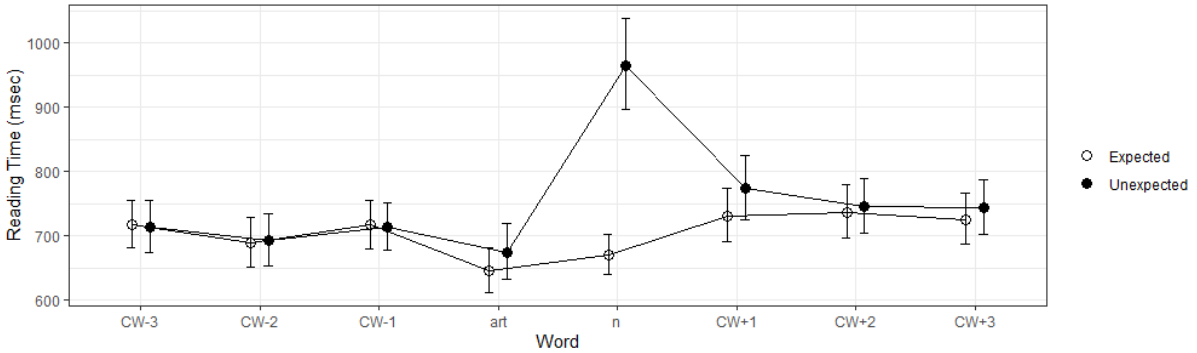


**Figure 1**: Reading times by region. Error bars represent 95% confidence intervals.

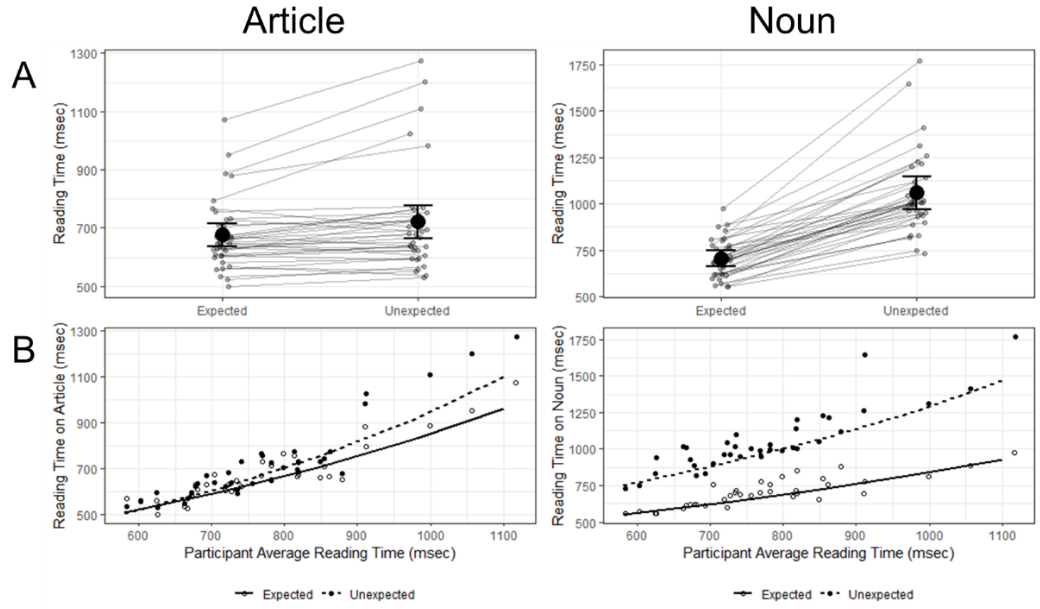


**Figure 2**: Reading times on the article and noun by expectation condition. (A) Participant average reading times and model fit reading times. (B) Participant average reading times on the article/noun against participant average reading times on the prior three regions. Lines indicate model fits.

**Selected References**. Boyce, Futrell, & Levy 2020 JML; Gulordava, Bojanowski, Grave, Linzen, & Baroni 2018 Proc. NAACL; Ito, Gambi, Pickering, Fuellenbach, & Husband 2020 Neuropsychologia; Nieuwland, et al, 2018 eLife; Witzel & Forster 2014 LCN; Witzel, Witzel, & Forster 2012 J Psyling Res.