

Expecting the unexpected: How discourse expectations can reverse predictability effects in reading time

Richard Futrell (Stanford University) & Hannah Rohde (University of Edinburgh)

futrell@stanford.edu

Discourse; Reading time; Mechanical Turk; Predictability

A growing body of work points to reduced processing difficulty when the words that listeners encounter are frequent or predictable given surrounding linguistic context (Kliegl et al., 2004; Levy 2008). Such work, however, fails to incorporate the notion that language is often used to convey new, and therefore unpredictable, information (Grice, 1975). If comprehenders expect content to be informative, one can ask whether certain contexts strengthen that expectation and, in so doing, make otherwise unpredictable material easier to process and conversely make highly predictable material difficult. While contextual support improves otherwise difficult-to-process constructions (Altmann & Steedman, 1988; Rohde et al., 2011; van Berkum et al., 1999), our study considers the opposite—whether context can increase the difficulty of material locally predictable from surrounding words, while rendering locally unpredictable (but highly informative) material easy.

In a self-paced moving-window reading-time study, we manipulated the local predictability of an instrument noun and comprehenders' global expectations regarding the informativity of upcoming material. A context sentence established global expectations for high or low informativity by characterizing an individual as either surprising or boring. The target sentence then described an activity that this individual performed with an instrument that was either highly inferable (1a,2a) or unpredictable (1b,2b). In keeping with established effects of predictability, inferable instruments were hypothesized to yield faster RTs than unpredictable instruments when readers expected low informativity (1a faster than 1b). However, when readers expected material to be informative, those same locally predictable instruments were hypothesized to be anomalously underinformative—and thus yield slower RTs (2a slower than 2b). Data was collected using Mechanical Turk, a “crowdsourcing” marketplace where workers are paid for small amounts of work. The advantage of a large crowdsourced population was that participants saw only a few items each and were unlikely to adapt to the frequent violation of their discourse expectations (#participants=110; mean #targets/participant=3.96).

The predicted interaction was observed at the final word of the sentence (mixed-effect model: $p < 0.001$). Pairwise comparisons showed unpredictable instruments being read *faster* than predictable instruments when readers expected high informativity ($p < 0.001$), whereas the opposite (numeric) pattern held when readers expected low informativity ($p = 0.12$). The fastest and slowest RTs overall were in the high-informativity condition: fastest for unpredictable instruments (*brush his teeth... knife*), slowest for predictable instruments (*chop some carrots... knife*). The predicted interaction appears numerically at the critical word. At the first spillover, predictable instruments were read faster than unpredictable instruments but only in the low-informativity condition, replicating existing effects of predictability, at least in discourse contexts that do not create strong biases for new and informative content.

These results echo evidence of listeners' sensitivity to cues signaling upcoming linguistic complexity or new information (Arnold et al., 2007; Jaeger, 2010) and work on production showing that speakers tend to mention only uninferable, unpredictable instruments (Brown & Dell, 1987). But to our knowledge these new findings are the first evidence of comprehension difficulty for material that is overly predictable from local cues. The results point to the importance of modeling comprehenders' pragmatic expectations about upcoming material—namely, their expectations about relevance and informativity.

(1) Sentence1: low-informativity expectation

My classmate Matthew is a boring person who always does things the way you'd expect.

a. Sentence2: locally predictable instrument [“don't expect surprise, don't get surprise”]

For instance, in order to chop some carrots, he was using a knife yesterday in the afternoon.

b. Sentence2: locally unpredictable instrument [“don't expect surprise but get surprise”]

For instance, in order to brush his teeth, he was using a knife yesterday in the afternoon.

(2) Sentence1: high-informativity expectation

My classmate Matthew is a surprising person who never does things the way you'd expect.

a. Sentence2: locally predictable instrument [“expect surprise but don't get surprise”]

For instance, in order to chop some carrots, he was using a knife yesterday in the afternoon.

b. Sentence2: locally unpredictable instrument [“expect surprise, get surprise”]

For instance, in order to brush his teeth, he was using a knife yesterday in the afternoon.