Online representations of implausible non-canonical sentences are more than good-enough

Michael G. Cutter (University of Nottingham), Kevin B. Paterson (University of Leicester), Ruth Filik (University of Nottingham).

Given an implausible non-canonical sentence such as '*The dog was bitten by the man'* people often state that the dog did the biting and the man was bitten, despite the opposite being true [1, 2]. This does not occur for the sentence in canonical form (*The man bit the dog*), suggesting that when an algorithmic parse is more complex readers may form 'good-enough' representations of sentences based on word order and pragmatic heuristics rather than an algorithmic parse. However, recent work suggests these findings may be attributable to task demands, with the paradigm used in [1,2] causing participants to query their sentence representation in a way that leads to misinterpretation [3,4]. We therefore tested whether we could find evidence for "good-enough" processing under conditions that did not impose explicit demands on comprehension.

We presented readers with two-sentence texts. The first sentence was implausible and shown in canonical (e.g. *It was the peasant that executed the king*) or non-canonical form (e.g. *It was the king that was executed by the peasant*). The following sentence was either 1) Algorithmically Consistent, such that it was plausible given a correct interpretation of the first sentence, but implausible given a good-enough interpretation of the first sentence (e.g. *Afterwards, the peasant <u>rode back to</u> the countryside*; the peasant is dead in a good-enough representation), or 2) Good-Enough Consistent, where the opposite was true (e.g. *Afterwards, the king <u>rode back to</u> his castle*; the king is dead in an algorithmic representation). If a good-enough representation is assigned to non-canonical sentences, we would predict an interaction between first sentence canonicality and follow-up sentence type. Specifically, reading times at the underlined region in the example sentences would be longer for the Algorithmically Consistent follow-up after a non-canonical than canonical sentence, and shorter for the Good-Enough Consistent follow-up after a non-canonical than canonical sentence. Given that prior work suggests older adults depend more on good-enough processing [5] we also compared effects for older vs. young adults.

We presented 44 items, normed for plausibility, with 80 filler items, to 120 participants (60 aged 18-25 years; 60 aged 65+ years) in non-cumulative phrase-by-phrase self-paced reading (see Fig. 1) using Gorilla.sc, a browser-based research platform [6]. We analysed log-transformed reading times for a target region at which implausibility emerged and a post-target region (see Fig. 1). Canonicality, Follow-Up Type, Age Group, and their interactions were set as predictor variables in a Bayesian mixed model. This showed no interaction of Canonicality and Follow-Up Type as a two-way interaction (see Fig. 2), or part of a three-way interaction with Age Group. A Bayes Factor analysis using default Cauchy priors favoured a null Canonicality * Follow-Up Type interaction (Target: BF_{10}=0.04; Post-Target: BF_{10}=0.05); however, there was strong evidence of longer reading times for Good-Enough Consistent vs. Algorithmically Consistent follow-ups in the Post-Target region (BF_{10} > 1000), suggesting that reading of the follow-up sentence was affected by its compatibility with the correct interpretation of the first sentence, with no evidence of misinterpretation. Young adults read faster than older adults (BF_{10} > 1000).

The results offer no evidence that participants formed "good-enough" representations of our sentences, rather than performing a full algorithmic parse. We argue that in the absence of specific task demands, participants do not arrive at a semantically incorrect interpretation of non-canonical sentences, consistent with the arguments put forward by [3].

C-AC: It was the peasant | that executed | the king.| Afterwards, | the peasant | rode back to | the countryside.
NC-AC: It was the king | that was executed by | the peasant. | Afterwards, | the peasant | rode back to | the countryside.
C-GEC: It was the peasant | that executed | the king. | Afterwards, | the king | rode back to | his castle.
NC-GEC: It was the king | that was executed by | the peasant. | Afterwards, | the king | rode back to | his castle.

Figure 1. An example of an item in each of our four conditions. "|" symbols represent the gaps between self-paced reading regions. "C-" and "NC-" represent canonical and non-canonical initial sentences, respectively. "-AC" represents a second sentence which is only plausible with an algorithmic parse of the first sentence, while "-GEC" represents a second sentence which is only plausible with a good-enough interpretation of the first sentence. The implausibility of the second sentence was always located in the third region of this sentence (e.g. *rode back to*), with a final sentence wrap-up region following this (e.g. *the countryside*/*his castle*).

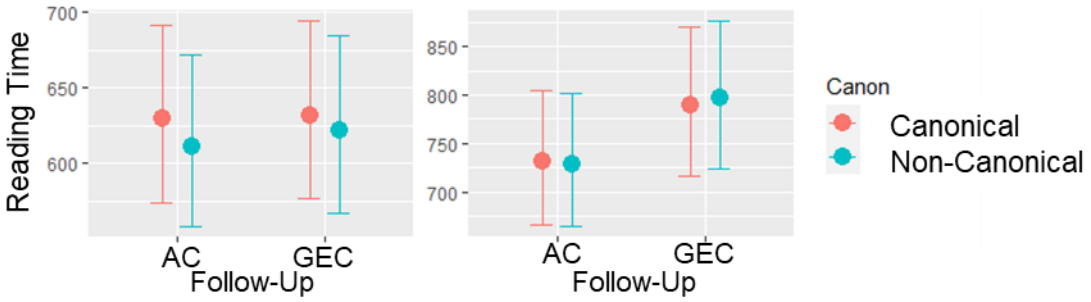

Figure 2. Predicted reading times from our Bayesian mixed models for our target region (left; *rode back to*) and post-target region (right; *the countryside*/*his castle*). AC represents the Algorithmically Consistent follow-up sentences, and GEC the Good-Enough Consistent follow-up sentences.

References: [1] Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203. https://doi.org/10.1016/S0010-0285(03)00005-7 [2] Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 538-544 https://doi.org/10.1037/a0018027 [3] Bader, M., & Meng, M. (2018). The misinterpretation of noncanonical sentences revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1286-1311. https://doi.org/10.1037/xlm0000519 [4] Meng, M., & Bader, M. (2021). Does comprehension (sometimes) go wrong for noncanonical sentences? *Quarterly Journal of Experimental Psychology*, *74*, 1-28. https://doi.org/10.1177/1747021820947940 [5] Christianson, K., Williams, C. C., Zacks, R. T., Ferreira, F. (2006). Younger and older adults' "good-enough" interpretations of garden-path sentences. *Discourse Processes*, *42*, 205-238. https://doi.org/10.1207/s15326950dp4202_6 [6] Anwyl-Irvine, A. L., Massonié, J., Flitton, A., Kirkham, N. & Evershed, J. K. (2020). Gorilla in our midst: An online behavioural experiment builder. *Behavior Research Methods*, *52*, 388-407. https://doi.org/10.3758/s13428-019-01237-x