Evidence for integration of noisy linguistic evidence and prior expectations depends on the task

Eli Kane & L. Robert Slevc (University of Maryland, College Park)
Eli.mikane@gmail.com

Psycholinguistic work on sentence processing has traditionally assumed "clean" error-free input to the processor and parsing has often been viewed as a modular process that is unaffected by non-syntactic factors such as context or plausibility (e.g., Frazier & Fodor, 1978). However, other work suggests that parsing and interpretation rely on rational integration of "noisy" language input (i.e., prone to error in transmission) with the likelihood of that perceived input (e.g., Levy 2008). An influential set of experiments that support this idea of a *noisy channel* parser comes from Gibson et al. (2013), who presented participants with sentences such as (1), which crossed plausibility and syntactic structure. Participants were likely to answer questions like (2) based on plausibility rather than literal syntax for implausible sentences, especially for sentences like (1d), where one need only assume a deletion (i.e., having not perceived an intended "for") compared to (1c) where one need assume an insertion (i.e., having perceived a non-intended "for").

(1) a. Plausible/PO: *The cook baked a cake for Lucy.*
     b. Implausible/PO: *The cook baked Lucy for a cake.*
     c. Plausible/DO: *The cook baked Lucy a cake.*
     d. Implausible/DO: *The cook baked a cake Lucy.*

(2) *Did the cake receive something?*

Because this research assessed listeners' interpretation via questions, the integration of plausibility with noisy input may not have occurred during parsing/interpretation of the sentence itself, but instead when evaluating the plausibility of answers to the following questions. To assess this possibility, we conducted two pre-registered experiments: first to replicate the effects shown by Gibson et al., and then to evaluate these effects when assessing listeners' sentence interpretations without asking questions (cf. Patson et al., 2009). Statistical power was estimated using simulations (Green & MacLeod, 2016) based on data from Gibson et al. (2013) and evidence for or against effects was assessed with Inclusion Bayes Factors (BF) on matched models using default priors. In both experiments, participants heard plausible and implausible variants of datives as in (1). In Experiment 1, following Gibson et al., participants answered a yes/no question following each sentence as in (2), where a non-literal interpretation was indicated by an incorrect answer to the question. We replicated those previous results: "literal" responses (based on the syntax) were less common for implausible than plausible sentences, especially for DO datives (Fig. 1; interaction $BF = 4.90$), suggesting that listeners integrated plausibility with the likelihood that the input parse was corrupted by noise.

Experiment 2 asked to what extent this integration of plausibility with the perceived input structure reflects processing of the parse/interpretation of the initial sentence versus processing of the question. Participants heard the same sentences, but instead of responding to yes/no questions, were simply instructed to type the sentence they heard. Here, a non-literal interpretation was indicated by a repetition using the opposite dative structure. Plausibility influenced repetition ($BF = 6,790$), but there was no evidence that this interacted with structure (Fig 2; interaction $BF = 1.00$). A cross-experiment comparison yielded strong evidence that the effect of plausibility differs across tasks (a plausibility-by-experiment interaction $BF = 167,000$).

These data suggest that comprehenders may indeed integrate syntactic parses with prior plausibility, however reliance on plausibility is more likely when an implausible parse is highlighted, for example when evaluating a previous interpretation while answering a question.
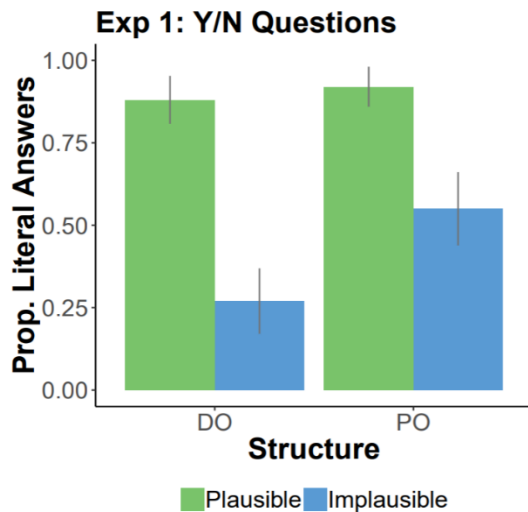
Figure 1: Proportion of yes/no question answers based on literal syntax as a function of plausibility and structure. Error bars indicate standard errors.
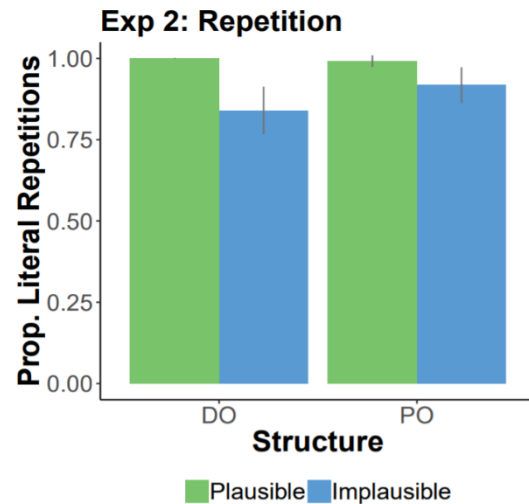

Figure 2: Proportion of sentence repetitions using the literal structure (compared to the alternate dative form) as a function of plausibility and structure. Error bars indicate standard errors.

References

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6,* 291-325.

Gibson, E., Bergen, L., Piantadosi, S.T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. PNAS, 110(20), 8051-8056.

Green, P. & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7,* 493-498.

Levy R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. *Proceedings of the 13th Conference on Empirical Methods in Natural Language Processing,* 234–243.

Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 280-285.