

## The Presentation Format in Cloze Tasks Influences Syntactic Predictability but It Does Not Influence Semantic Predictability

Bozhidara Hristova, Robin Lemke, Lisa Schäfer, Heiner Drenhaus & Ingo Reich  
(Saarland University)

bozhidara.hristova@uni-saarland.de

In research on predictability effects in language comprehension, predictability is often operationalized in terms of cloze probability. Although predictability effects based on cloze probabilities have been consistently attested in research (e.g., Rayner & Well 1996, Federmeier & Kutas 1999, Smith & Levy 2011), the concern has emerged that the standard cloze task (Taylor 1953) does not capture the influence of memory decay on predictability (*lossy-context surprisal*, Futrell et al. 2020). For this reason, Apurva and Husain (2021) perform cloze norming with a new paradigm, in which the context is presented in a self-paced reading (SPR) format. However, it remains unclear whether a new paradigm is needed, i.e. whether the presentation format in cloze experiments affects the obtained probabilities.

We investigate this question on the case of a cloze task conducted to test for predictability effects on the usage of Gapping (Exp. 1,  $N = 160$ ). Since the obtained cloze probabilities did not predict reading times in an SPR experiment in the expected direction, we hypothesized that they did not model the expectations of the participants in the SPR experiment accurately. We therefore conducted a written sentence completion study with a centered SPR presentation (Exp. 2,  $N = 48$ ) and compared the responses to the ones from Exp. 1. We focused on the predictability of the verb in the second conjunct (C2) of parallel coordinations (i.e. coordinations with the same verb in both conjuncts) (1) and intended to manipulate the predictability of the C2 verb through the number of objects in the context sentence (OBJECT NUMBER, one/two) (2). We hypothesized that mentioning only one object in the context would increase the probability of the C2 verb. In Exp. 1, the context/target pairs were displayed together with a text box after the C2 subject. In Exp. 2, subjects read the items word by word by pressing the spacebar. After reading the C2 subject, a text box appeared. In both studies, participants were asked to type in the continuation they considered most likely. We aimed to test whether the SPR format would change the produced continuations due to participants accessing the preceding linguistic context only from memory. For the probability of a parallel (same-verb) response, we expected that the SPR presentation would make our manipulation less effective, i.e. that OBJECT NUMBER will have a weaker effect on the probability of a parallel continuation in Exp. 2 than in Exp. 1 (OBJECT NUMBER  $\times$  PRESENTATION TYPE interaction). Regarding the syntactic realization of the parallel response (Gapping vs nonelliptical), we hypothesized that Gapping will be less frequent in Exp. 2 than in Exp. 1, since its licensing is conditioned on the content of the first conjunct (main effect of PRESENTATION TYPE).

With respect to the probability of a parallel response, we found no significant interaction between OBJECT NUMBER and PRESENTATION TYPE ( $z = -0.46$ ,  $p = 0.65$ ) and no main effect of PRESENTATION TYPE ( $z = -0.37$ ,  $p = 0.71$ ) (Fig. 1). The similarity between the two experiments was also evident in the overall uncertainty in the sample of responses, with no main effect of PRESENTATION TYPE ( $z = -0.17$ ,  $p = 0.87$ ) and no OBJECT NUMBER  $\times$  PRESENTATION TYPE interaction effect ( $z = 0.15$ ,  $p = 0.88$ ) on the entropy of the produced C2 verb per item (Fig. 2). However, looking at the form of the parallel continuations, we observed a significantly lower probability of Gapping in Exp. 2 than in Exp. 1 ( $z = -3.34$ ,  $p = < 0.001$ ) (Fig. 3). The results suggest that syntactic predictability is more sensitive to the accuracy of memory representations than the predictability of semantic content. One reason could be that producing an unlicensed syntactic structure will lead to an ungrammatical sentence. No such risk is associated

with producing a semantically different verb. Thus, when subjects are uncertain about the preceding context, they opt for the safe (nonelliptical) syntactic option more often.

(1) *Anna is knitting a sweater, and Max **(is knitting)** a scarf.*

(2) *Die Anna und der Max haben im Bastelladen (**Wolle** | **Wolle und Origamipapier**) the Anna and the Max have in craft.store wool wool and origami.paper gekauft. Die Anna strickt einen Pulli und der Max \_\_\_\_\_.*  
 bought the Anna knits a.ACC sweater and the.NOM Max \_\_\_\_\_

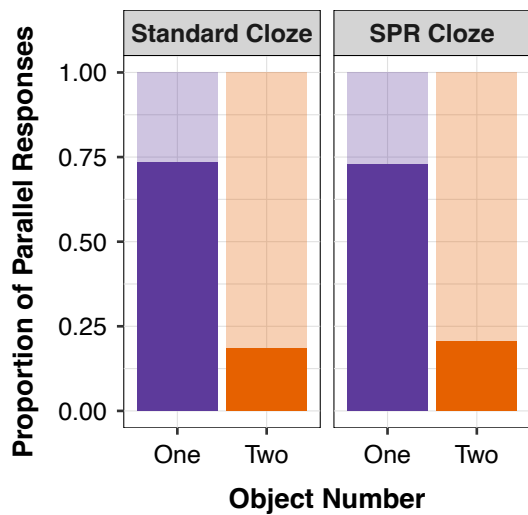


Fig. 1 Proportion of parallel (same-verb-as-in-first-conjunct) continuations

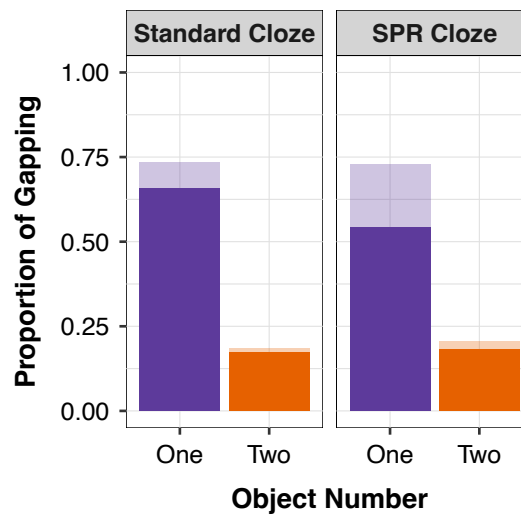


Fig. 3 Proportion of Gapping in the parallel responses

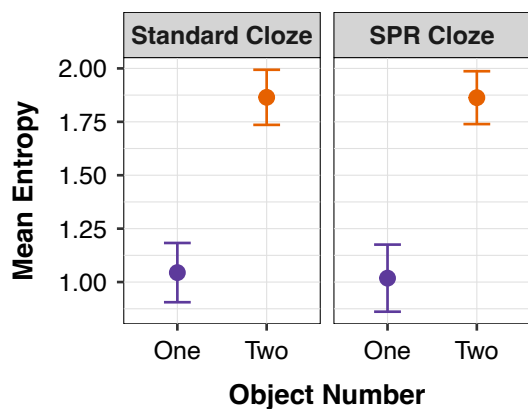


Fig. 2 Mean entropy and SE per item in the produced C2 verbs. For the calculations for the standard cloze task, we used the averaged entropy values from 4000 samples with the size of 48 subjects.

## References

- Apurva, & Husain, S. (2021). Revisiting anti-locality effects: Evidence against prediction-based accounts. *Journal of Memory and Language*, 121, 104280.
- Federmeier, K. D., & Kutas, M. (1999). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, 8(3), 373–392.
- Futrell, R., Gibson, E., & Levy, R. (2020). Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive Science*, 44, e12814.
- Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3(4), 504–509.
- Smith, N. J., & Levy, R. (2011). Cloze but no cigar: The complex relationship between cloze, corpus, and subjective probabilities in language processing. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 1637–1642). Cognitive Science Society.
- Taylor, W. L. (1953). “Cloze procedure”: A new tool for measuring readability. *Journalism Quarterly*, 30, 415–433.