

Predictability effects on the acceptability of antecedent-target mismatches under verb phrase ellipsis

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

We present an information-theoretic account of acceptability contrasts between antecedent-target mismatches under verb phrase ellipsis (VPE) and three experiments which support its prediction: The acceptability of mismatches (and VPE in general) increases as a function of the context-based likelihood of the target of ellipsis. Our approach relies on general processing mechanisms rather than specific assumptions about the syntax and processing of VPE mismatches, which have been proposed in previous research.

Since Sag (1976); Williams (1977), it has been assumed that verb phrase ellipsis (VPE) requires a syntactically identical antecedent, as in (1a), whereas structural differences like the voice mismatch in (1b) are not acceptable. However, (1c) shows that voice mismatches are not unacceptable across the board.

- (1) a. Ann invited Bill and Sue did ⟨invite Bill⟩, too.
b. *Ann invited Bill and Sue was ⟨invited by Ann⟩, too.
c. This problem was to have been looked into, but obviously nobody did ⟨look into the problem⟩. (Kehler, 2002, 548)

1 Background

Mismatches have been used as diagnostics in the ongoing debate on the syntactic analysis of VPE. On the one hand, *syntactic* accounts claim that the ellipsis site contains rich syntactic structure, which is deleted under identity with the antecedent (e.g. Sag, 1976). On the other hand, *pragmatic* accounts (e.g. Hardt, 1993) argue that VPE involves a null anaphor resolved through context. Unacceptable mismatches support syntactic accounts, because VPE without identity is degraded, whereas acceptable mismatches favor pragmatic accounts by suggesting that full identity is not necessary. From the syntactic perspective, which in principle predicts identity, acceptable mismatches have been addressed in two ways: They could be grammatical, but sometimes unacceptable (Kim et al., 2011), or they could be ungrammatical, but sometimes acceptable (Arregui et al., 2006; Parker, 2018). The mismatches between acceptability and grammaticality have been accounted for by different processing mechanisms: Kim et al. (2011) argue that unacceptable mismatches violate the parser's expectations. Arregui et al. (2006) propose a VPE-specific reconstruction mechanism, which allows hearers to use syntactic derivation to construct a matching antecedent if none is present. Parker (2018) proposes a more general noisy memory retrieval mechanism. While Kim et al. (2011) argues with specific parsing rules, the latter two accounts predict that, in general, mismatches are more acceptable when antecedent and target are syntactically similar to each other. A problem for these accounts is that their predictions rely on syntactic properties of the antecedent, the target, or the degree of syntactic similarity between these expressions. However, the acceptability of mismatches is also driven by nonsyntactic factors like discourse relation (Kehler, 2002), information structure (Kertz, 2013), implicit Questions under Discussion (QuD) (Miller and Hemforth, 2014) and even extralinguistic context (Geiger and Xiang, 2021). In sum, these studies indicate that VPE mismatches improve when the target is likely in context, for instance because it refers to a salient QuD or its information structure is aligned with the antecedent. If structurally similar conjuncts are relatively likely, likelihood might also play a role in the similarity-driven effects observed by Arregui et al. (2006).

2 Information-theoretic account

The reduction of predictable material is a crucial prediction of information-theoretic accounts of linguistic encoding, which derive it from a general tendency toward distributing the hearer's processing effort uniformly across the utterance. Since predictable expressions are easier to process (Hale, 2001), reducing them avoids underutilizing processing resources (Fenk-Oczlon, 1989; Levy and Jaeger, 2007). In turn, using longer forms for unpredictable expressions avoids exceeding the processing resources. Applied to VPE, we therefore expect that mismatches (but in principle any instance of VPE) is more strongly preferred when the target is more likely. This likelihood might be modulated by diverse factors, such as linguistic context, the form of the antecedent, pragmatic inferences, and extralinguistic context. Therefore, the information-theoretic account is potentially capable of explaining some of the previous empirical findings with a single and independently evidenced processing mechanism, which has been shown to constrain the distribution of other omission and reduction phenomena (Jaeger and Buz, 2017).

3 Experimental rationale

We test the predictions of the information-theoretic account at the case of the stimuli used by Arregui et al. (2006), which they provide in their appendix. Across the four conditions in (2), Arregui et al. (2006) find a gradual acceptability cline ($a > b > c > d$), which they interpret as indicating the effort of constructing a parallel antecedent by syntactic derivation: When antecedent and target are similar, this effort is low and the mismatch is relatively acceptable, but the more derivation steps are required to build a matching antecedent, the more unacceptable is the mismatch. We hypothesize that the acceptability cline might result from the decreasing likelihood of the target from (2a) through (2d), which our information-theoretic account predicts to result in higher processing effort and degraded acceptability of VPE. We use an acceptability rating experiment to replicate the pattern in Arregui et al. (2006), a production task to measure the likelihood of the target VP and corresponding production preferences and a self-paced reading study to test for effects on processing effort.

- (2) a. None of the astronomers saw the comet, but John did. (Available VP)
 b. Seeing the comet was nearly impossible, but John did. (Embedded VP)
 c. The comet was nearly impossible to see, but John did. (VP with trace)
 d. The comet was nearly unseeable, but John did. (Negative adjective)

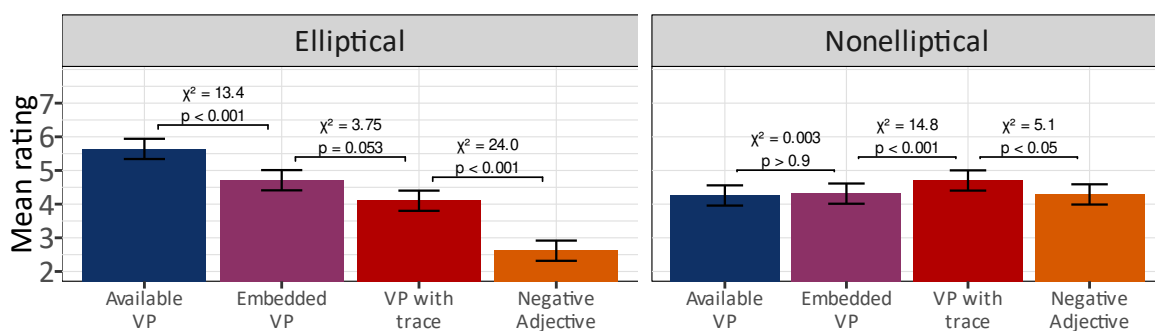


Figure 1: Mean acceptability ratings for VPE (left) and nonelliptical controls (right).

4 Experiment 1 – Acceptability rating

We first replicated the data by Arregui et al. (2006) in a web-based rating study conducted with LimeSurvey (LimeSurvey GmbH, 2012). Unlike Arregui et al. (2006), we also included the corresponding nonelliptical utterances (...but John saw the comet.) to investigate whether the effect is specific to ellipsis. Arregui et al. (2006) tested the first conjuncts as controls, but this does not rule out the possibility of an ellipsis-independent mismatch penalty (Kim et al., 2011).

The 16 items were mixed with 60 fillers and rated on a 7-point Likert scale (7 = fully acceptable) by 96 subjects recruited on the *Clickworker* crowdsourcing platform. COMPLETENESS was tested between subjects to ensure that the results are comparable to Arregui et al. (2006). The data (see Fig 1) were analyzed separately for the elliptical and nonelliptical conditions with cumulative link mixed models for ordinal data (Christensen, 2019). The predictor CONSTRUCTION was forward-coded, so that each of the three contrasts compared a level to the subsequent one(s). For the elliptical conditions, this replicates the pattern reported by Arregui et al. (2006). The nonelliptical sentences show that the effect is specific to ellipsis: The contrasts are either not significant, or their direction is inverted.

5 Experiment 2 – Production

We used a web-based written production task implemented in LimeSurvey to investigate two predictions of the information-theoretic account: (i) The potentially omitted VP is more likely in conditions where it is more acceptable and (ii) subjects produce more instances of VPE in that case. 120 participants recruited on *Clickworker* were asked to provide the most natural continuation of the materials, which were cut off after the subject in the 2nd conjunct (*John, in (2)*). The data were annotated for whether subjects produced a VP identical to the one that would be omitted under VPE (*saw the comet*) and, if so, whether this VP was reduced by VPE. The data (Fig. 2) were analyzed with logistic mixed effect regressions (Bates et al., 2015) predicting one of the binary dependent variables IDENTITY or ELLIPSIS from the forward coded CONSTRUCTION predictors. The analyses show that in two out of three contrasts, the VP is significantly more often produced in conditions where VPE is judged as more acceptable in the rating study, and that – among the VPs that can be reduced – VPE is more frequent the more likely the VP is. Taken together, this shows that the acceptability differences in experiment 1 are related to a gradual decrease in likelihood from (2a) through (2d), which is also reflected in a stronger preference for omitting the VP in a production task.

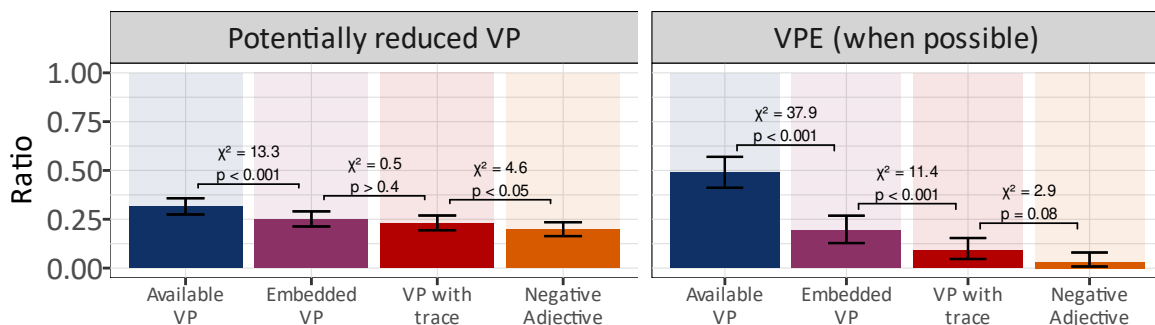


Figure 2: Ratio of potentially reduced VPs (left) and VPE ratio among these VPs (right).

6 Experiment 3 – Self-paced reading

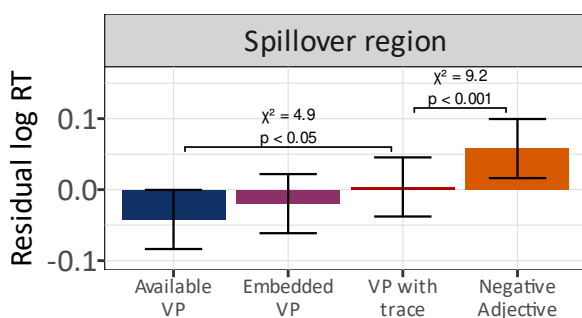


Figure 3: Residual log RTs at the SOR onset.

We then used a web-based self-paced reading study to investigate whether the link between the target VP likelihood and the preference for VPE is mediated by processing effort. 48 subjects recruited on *Clickworker* read 16 items like (2), which were extended with a spillover region (SOR, a causal or temporal clause like *because he had a special telescope* for (2)) and mixed with 60 fillers. The stimuli were presented word-by-word in

a centered self-paced reading paradigm using PCIBex (Zehr and Schwarz, 2018). We analyzed the mean residual log reading times of the first three words of the spillover region which followed the auxiliary *did* (Fig. 3) with linear mixed effects regressions (Bates et al., 2015) using the same predictors as in the previous experiments. The analysis reveals a similar gradual pattern as experiments 1 and 2: Ellipsis seems to be more difficult to process when the VP is less likely. In the main analysis (forward coding), only one of the CONSTRUCTION contrasts is significant (VP with trace vs. negative adjective). However, pairwise comparisons show that the difference between the available VP and VP with trace conditions is also significant (due to the gradual difference between the first three conditions, forward coding cannot test this).

7 Discussion

Our experiments show that the gradual acceptability pattern reported by Arregui et al. (2006) is in line with production preferences and processing effort: As the information-theoretic account predicts, VPE is more acceptable when the omitted VP is more likely (in this case, subjects are also more likely to produce VPE) and it is more easily processed. The data do not speak against Kim et al. (2011) and Kertz (2013), but the predictors they investigate do not explain the complete pattern in our data. Our findings are partially in line with previous accounts of mismatches, in particular with Arregui et al. (2006) and Parker (2018). However, Arregui et al. (2006) assume VPE-specific repair mechanism, whereas our account operates on more general processing principles. Furthermore, Arregui et al. (2006) and Parker (2018) operate on the similarity between antecedent and target, so that they cannot take into account pragmatic or extralinguistic factors, which have been shown to modulate the acceptability of mismatches in other studies. The information-theoretic account predicts effects of all factors which modulate the likelihood of the target. A further implication is that some of the identity conditions proposed in the literature for VPE and other ellipses might be traced back to differences in predictability. To what extent this is possible must remain open for future research, and since information-theoretic optimization is limited to a choice between grammatical structures (Jaeger, 2010), some mismatches might still be ruled out by grammar. Taken together, the information-theoretic account is supported by the data, it provides a uniform explanation of VPE mismatches and other omissions, and it is based upon independently motivated processing principles.

Selected references

- Arregui, A., Clifton, C., Frazier, L., and Moulton, K. (2006). Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language*, 55(2): 232–246.
- Geiger, J. and Xiang, M. (2021). At the syntax-discourse interface: Verb phrase ellipsis interpretation in context. *Language*, 97(1):e89–e110.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of NAACL (Vol. 2)*, pages 159–166.
- Hardt, D. (1993). *Verb phrase ellipsis: Form, meaning, and processing*. Number 182 in IRCS Technical Reports Series.
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. CSLI, Stanford, California.
- Kertz, L. (2013). Verb phrase ellipsis: The view from information structure. *Language*, 89(3):390–428.
- Kim, C. S., Kobele, G. M., Runner, J. T., and Hale, J. T. (2011). The Acceptability Cline in VP Ellipsis. *Syntax*, 14(4):318–354.
- Levy, R. P. and Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In Schölkopf, B., Platt, J., and Hoffman, T., editors, *Advances in neural information processing*, volume 19, pages 849–856. MIT Press, Cambridge, MA.
- Merchant, J. (2013). Voice and ellipsis. *Linguistic Inquiry*, 44(1):77–108.
- Miller, P. and Hemforth, B. (2014). *VP Ellipsis beyond syntactic identity: the case of nominal antecedents*.
- Parker, D. (2018). A memory-based explanation of antecedent-ellipsis mismatches: New insights from computational modeling. *Glossa: a journal of general linguistics*, 3(1):129.
- Sag, I. A. (1976). Deletion and logical form. PhD thesis, MIT, Cambridge, MA.
- Williams, E. (1977). Discourse and logical form. *Linguistic Inquiry*, 8:101–139.