Discourse expectations guide the use of verb phrase ellipsis: Psycholinguistic evidence for an information-theoretic account of context and length effects on VPE

Lisa Schäfer, Robin Lemke, Heiner Drenhaus & Ingo Reich (Saarland University)
lisa.schaefer@uni-saarland.de

Overview. Verb phrase ellipsis (VPE) (1a) is probably one of the most extensively studied omission phenomena in English. But both the theoretical and the psycholinguistic literature have focused on its grammatical properties, leaving aside *why* speakers use ellipsis (1a) instead of a full form (1b) at all.

(1) a. Sam played football in the backyard of the house and Dean did, too. (*ellipsis*)
b. Sam *played football in the backyard of the house* and Dean *played football in the backyard of the house*, too. (*full form*)

We pursue the hypothesis that the use of VPE by a speaker is driven by discourse expectations that lead to audience design (Bell 1984): The speaker omits parts of an utterance that she thinks are highly predictable and that would thus be redundant to the listener.

Account. We model this strategy with the information-theoretic concept of Uniform Information Density (UID) (Levy & Jaeger, 2007). According to UID, speakers tend towards distributing information ($-log_2\ p\ (word|context)$ (Shannon 1948)) as uniformly as possible across utterances, avoiding troughs in the information density (ID) profile. Such troughs are regions with continuously low, i.e. redundant, information. As information indexes processing effort (Levy 2008) on the part of the listener, such troughs caused by redundancy hamper comprehension and should be avoided by the speaker.

(2) a. Sam and Dean dream of becoming NFL quarterbacks some day. (*predictive context*)
b. Sam and Dean dream of becoming President some day. (*unpredictive context*)

The redundancy of an utterance could be modulated by discourse context: In a predictive context like (2a), Dean should be more likely to also play football than in the unpredictive context of (2b). This should lead to a higher pressure on the speaker to omit the redundant VP and to utter (1b) in the context of (2a) than in the context of (2b) as the trough created by the redundant repetition of the VP is deeper in the predictive than in the unpredictive context (Fig. 1).

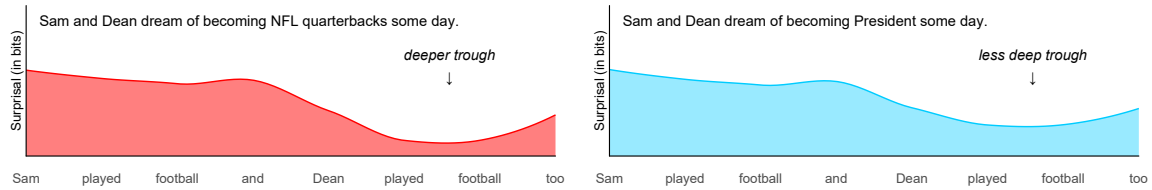

Fig. 1: Hypothetical information density profiles of the example in unpredictive and predictive context

UID predicts not only that utterances containing deeper troughs are less well-formed, but also those containing longer troughs. A trough is longer the longer the redundant part of an utterance is. As a consequence, the pressure to use VPE should be greater for a redundant long VP like *play football in the backyard of the house* than for a redundant short VP like *play football*. For both context and length effects, we obtain preferences for VPE with rating studies and correlate them with cognitive effort as indexed by reading times which we measure using self-paced reading (SPR).

Context – Pretest. We pretested our contexts to ensure that the event denoted in the target sentence was significantly more likely to occur in the predictive context (2a) than in the unpredictive context (2b). 48 participants rated this likelihood on a scale from 0 (cannot happen) to 100 (must happen) yielding a numeric ContextScore for each item and condition. We included this score in our analysis instead of a binary variable to capture the differences between items.

Context – Rating. To test whether listeners indeed prefer VPE more in the predictable condition, we conducted an acceptability rating study with 95 native speakers of English recruited via *Prolific Academic* in a 2 × 2 (Context: predictive vs. unpredictive × Completeness: ellipsis vs. full form) design. Each participant judged 12 of 24 items (6 in a predictive (2a) and 6 in an unpredictive context (2b)) and 84 fillers on a 7-point Likert scale (7 = completely natural). The manipulation of Completeness was done between subjects. We analyzed the data with Cumulative Link Mixed Models (package `ordinal` (Christensen 2018), R) with the full random effects structure justified by the data (Barret al. 2013). The analysis revealed a significant main effect ($\chi^2 = 35.12$, $p < .001$)of ContextScore that indicates that

participants rate utterances as better when they occur in a predictive context. A significant interaction (χ^2 = 6.1, p < .05) between COMPLETENESS and CONTEXTSCORE shows that this preference is particularly strong in the elliptical condition (Fig. 2). This result supports our hypothesis: The effect of predictability is stronger for elliptical than for non-elliptical utterances.

CONTEXT – SPR. Surprisingly, a newly conducted self-paced reading task on only the full forms (1b) did not reveal a significant difference in processing effort indexed by reading times of the second conjunct between the predictive and the unpredictive condition. In our presentation, we discuss possible reasons, for example whether the parallelism effect induced by the identical first conjunct is strong enough to overwrite the context effect.

LENGTH – Rating. We conducted a rating study with 41 subjects, 32 items like (1) without context and 72 fillers in a 2 × 2 (FORM: ellipsis vs. full form × LENGTH: short vs. long) within-subjects design following the procedure of the context rating study. In the short condition we presented a plain VP (*played football*), which was expanded by a PP in the long condition (1). We expected a relatively higher preference for VPE in the long condition due to the longer trough in the ID profile. The analysis with CLMMs with full random effects structure revealed a significant main effect of LENGTH (χ^2 = 27.92, p < .001) showing that overall items with longer VPs were rated worse (Fig. 3). A significant FORM:LENGTH interaction (χ^2 = 7.66, p < .01) indicates that the full form was rated particularly worse in the long condition. This is in line with our hypothesis: Listeners prefer VPE more strongly when the repetition of the antecedent VP is more redundant.

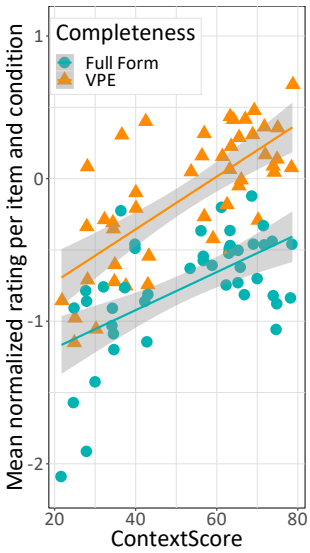

Fig. 2: Normalized ratings as function of CONTEXTSCORE

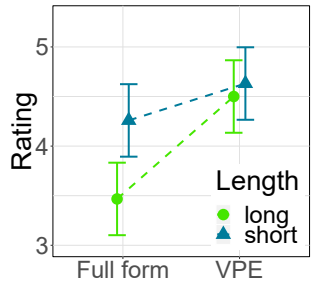

Fig. 3: Mean ratings and 95% CIs for LENGTH

LENGTH – SPR Rating data however do not provide direct evidence for troughs in the ID profile. We used a SPR study on the full forms to measure the processing effort on the redundant VP. We expected to find a longer trough in the more redundant long condition, which would be indicated by a larger difference in average reading time between the 1st and the 2nd conjunct in the long condition. Using IBEX (Drummond 2016), the items and fillers from the rating study, extended by a spillover region, were presented word-by-word centered to 96 participants. Our DV were cumulative reading times per CONJUNCT (italicized in (1b)) which we residualized per subject (Gibson & Levy 2016). Linear mixed effect models (`lme4` (Bates et al. 2015)) with a random intercept for items and by-subject and by-item random slopes for LENGTH and CONJUNCT revealed a significant main effect of LENGTH (χ^2 = 13.11, p < .001) and of CONJUNCT (χ^2 = 37.12, p < .001) showing that generally short items were read faster than long ones and that second conjuncts were read faster than first ones. A significant interaction between LENGTH and CONJUNCT (χ^2 = 75.18, p < .001) indicates that the second conjunct was read especially faster in the long condition (Fig. 4). This supports our hypothesis: The longer VP is more redundant and creates a longer trough, which is reflected in faster reading times.

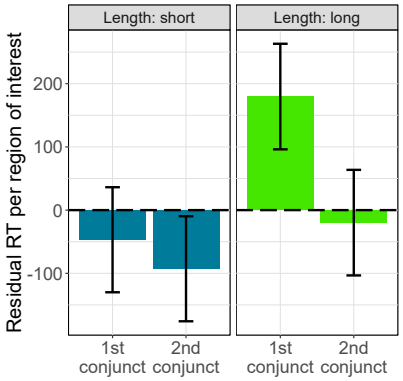

Fig. 4: Cumulative residual RT per region of interest and 95 % CIs for SPR

Discussion. Our data, except for the results from the CONTEXT SPR that still need explaining, provide further evidence for the effect of UID on encoding preferences in line with our hypothesis: Speakers omit redundant parts of the utterance and use ellipsis to avoid troughs in the ID profile. The preference for omission increases in predictive contexts and for longer redundant elements. Speakers anticipate the discourse expectations of the listener and perform audience design.

Selected references. Bell, A. (1984): Language Style as Audience Design. In: *Language in Society* 13(2): 145–204. • Levy, R. (2008). Expectation-Based Syntactic Comprehension. In: *Cognition* 106(3): 1126–1177. • Levy, R. P. and Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In Schlökopf, B., Platt, J. C., and Hoffman, T., editors, *Advances in Neural Information Processing Systems* 19: 849–856.