How far ahead does prediction span? The timing of verb pre-activation during sentence comprehension

Jeonghwa Cho[1], Shota Momma[2], Jonathan R. Brennan[1]
1.University of Michigan, 2. University of Massachusetts

Prediction has been proposed as a core component of language comprehension, but previous studies often focused on immediately preceding words (e.g., [1], [2]). This study examines how far ahead a verb may be pre-activated. We focus on contexts where the verb is highly predictable and test whether structural properties of the target verb modulate the timing of verb prediction. We compare unaccusative and transitive sentences, which have shown different timing of verb selection during production ([3]) such that unaccusative, but not transitive, verbs are selected before the subject noun. This is taken to follow from the status of the subject as an internal argument for unaccusative verbs. If the timing of verb prediction during comprehension is also affected by argument-structure properties, we expect to find earlier verb activation for unaccusative sentences than transitive sentences. To test this we use representational similarity analysis following Hubbard and Federmeier [4] by computing the spatial correlation between EEG data of the target verb and words earlier in the same sentence. **METHODS** EEG data from 32 English speakers were collected (32 channels, 500 Hz) while participants read 72 English scenarios consisting of two sentences (See Table 1). The latter part of the second sentence always had the structure *DET–ADJ–N–ADV–V(–DET– N*), where the target verb is unaccusative (*n* = 36) or transitive (*n* = 36), with equal number of animate and inanimate subjects for each verb type. Cloze probability norms (N=30 English speakers) confirmed high predictability of verbs in both types (unaccusative M=84.89%, SD=12.17, transitive M=80.00%, SD=11.99). The ADJ-N pair was repeated across the two sentences to ensure that the expected subject is not ambiguous. Each scenario was presented word by word for 500 ms followed by a blank screen of 300 ms (Fig.1), which participants read silently and answered comprehension questions for 10% of the trials. EEG data were epoched between -100 – 800 ms from the onset of *DET, ADJ, N, ADV,* and *V*. Pearson's *r* was computed across sensors at each time point between verb epochs and epochs from each preceding word. **RESULTS** Figure 2A summarizes correlations between words across time. Differences between Word Position and Verb Type were assessed with a two-factor within-subjects ANOVA on data averaged within 160-210ms (Fig 2B, time-window from [4]). There was a reliable effect of Word Position and, crucially, an interaction with Verb Type at ADJ (F=4.01, p=0.048) such that correlations at ADJ were significantly different from those at DET for unaccusative (F=5.52, p=0.004) but not for transitive (F=0.31, p=0.578) verbs. There was no effect of Verb Type at the N or ADV, and the correlations at both words were greater than DET (ps < 0.004). Semantic similarity (GloVe) and surprisal (GPT-2) between the adjective and the verb were computed to check whether the correlation is a mere reflection of these measures. If this the case, those measures would be different by Verb Type at the adjective. However, pairwise comparisons showed this is not the case (ps > 0.369). **DISCUSSION** In summary, we observe statistically significant differences in the correlation between EEG signals earlier in the time window of 160-210 ms, which has been identified as key for next word prediction in this paradigm [4]. Expanding previous results, we find verb pre-activation not only at the immediate previous word (i.e., adverb) but also at earlier words (i.e., subject noun and even adjective for unaccusative verbs) when the verb is highly predictable. In parallel to sentence production ([3]), this pre-activation is sensitive to the argument structure that is not captured by surprisal or word similarity. As the sentences in this study were highly constrained, further research on more naturalistic settings may be beneficial.

Table 1. Examples of experimental stimuli

Verb type	Scenario
Unaccusative	(a) A beautiful snowman was standing on a warm sunny day. When the children came back from school, they found that **the beautiful snowman finally <u>melted</u>.** *DET ADJ N ADV V*
Transitive	(b) On a windy winter day, a massive fire broke out behind the village. People who were going about their daily lives fled, and **the massive fire eventually <u>burned</u> the village.** *DET ADJ N ADV V DET N*

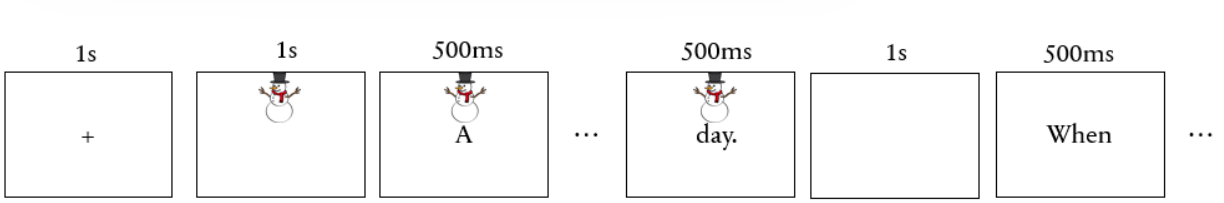

Figure 1. Example trial for the scenario "(a) A beautiful snowman was standing on a warm sunny day. When the children came back from school, they found that **the beautiful snowman finally melted."** The image of the noun was presented along with the first sentence to facilitate pre-activation of the subject in the following sentence, and therefore mitigating against uncertainty of the subject being a confounding factor.

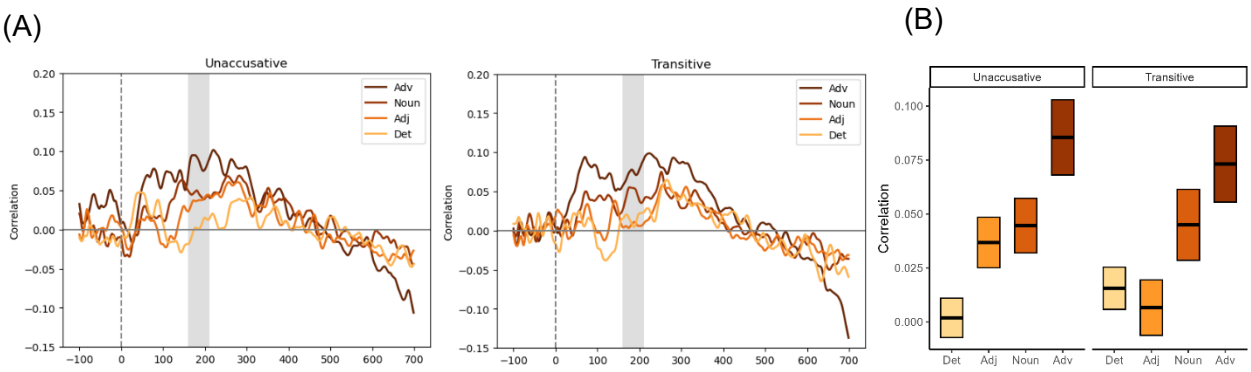

Figure 2. (A) Correlation between each word and the final verb for unaccusative and transitive sentences. (B) Mean correlation between each word and the final verb for unaccusative and transitive sentences in the time window of 160 – 210 ms (shaded in Figures 2A). Upper and lower limits indicate standard errors.

Selected references [1] DeLong et al., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. [2] Kutas and Hillyard, 1984. Brain potentials during reading reflect word expectancy and semantic association. [3] Momma and Ferreira, 2019. Beyond linear order: The role of argument structure in speaking. [4] Hubbard and Federmeier, 2021. Representational pattern similarity of electrical brain activity reveals rapid and specific prediction during language comprehension.