

## Online cloze evidence for rapid use of lexical and grammatical cues

Masato Nakamura & Colin Phillips (University of Maryland)

Predictions about upcoming input are standardly measured via facilitated processing of explicitly presented words (fixation times, N400 amplitudes) or anticipatory looks in scenes (e.g. [1, 2]). In this study we examine predictions via a spoken, speeded cloze task in Japanese. We use information from spoken responses to understand how and when contextual cues are used to generate predictions, revealing effects obscured in EEG studies.

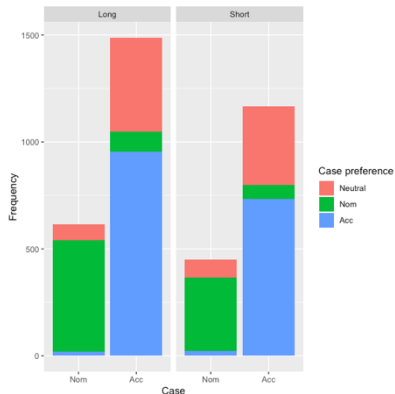
Situations where online measures of prediction diverge from corpus probabilities or late measures are particularly informative about how predictions arise. A useful test case is argument role reversals, in which an anomalous verb is processed as if it is more expected than it should be (e.g. [3,4]). For example, in *The customer that the waitress served* vs. *The waitress that the customer served* the verb *serve* differs in offline cloze probability, but EEG studies in many languages have found that it elicits identical N400 amplitudes. Additional time between the arguments and the verb yields an N400 contrast [5]. These findings motivated the claim that early predictions reflect lexical associations, with role-specific predictions emerging only after a delay. However, the explicit presentation of anomalous verbs in these studies might bias the estimate of how expected those anomalous words were.

We examined the timing of use of argument role and lexical cues in a Japanese speeded cloze task, using materials from a previous EEG study [6]. We presented minimal contexts of a noun and a case marker, which participants completed with a verb. The cloze task measures predictions via speakers' own productions. Instead of measuring the degree of convergence of open-ended predictions (i.e. cloze probability), we used the full set of productions to test predictions at specific times by (i) using simple contexts, to control lexical and grammatical content of cues, (ii) limiting the response time windows [7] and (iii) using a simple NLP measure to assess the relationship between contexts and produced items. This was possible by gathering spoken responses via the internet.

80 speakers [40 analyzed so far] each completed 160 visually presented fragments. In a 'long' block responses had to start after 1.6-2.8s, and by 1.2s in a 'short' block. The timeline and the stimuli matched an existing study that found identical N400s at the verb, regardless of case. For each of the 5389 produced noun-verb pairings we measured speech onset latency, noun-verb similarity using Japanese word2vec [8], and whether the pairing would be more plausible with nominative or accusative case-marking, e.g., *thief-acc arrest* is more plausible than *thief-nom arrest*. Pairings featuring the dispreferred case were coded as reversals.

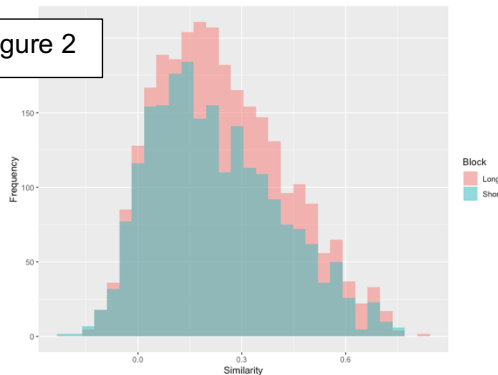
Argument roles clearly affected productions. Case-compatible productions were far more common than reversals, comprising 94.4% of trials in the short condition (Fig. 1). Verb transitivity clearly matched the case marking. Noun-verb similarity was higher in the long condition, suggesting more specific expectations with more time (Fig. 2). The verbs produced in reversed responses tend to have high cloze probabilities in the other case markings, suggesting that role-independent lexical associations serve as lures (Fig 3.). Speech onset latencies were shorter for more similar pairings. Overall, the speeded cloze results show that both argument roles and lexical association shape early predictions [cf. 9]. The discrepancy with prior EEG results could reflect a monitoring process that filters (most) role-incompatible productions in the cloze task, or a biasing effect of explicitly presented lures in EEG studies.

Figure 1



The frequency of transitive verb productions. The x-axis represents the case marking of the context noun, and the color represents whether that verb is more plausible if that noun is in nominative or accusative case, or is neutral. Very few noun-verb productions involved the dispreferred case marker

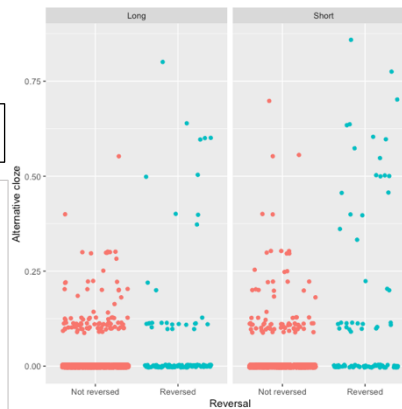
Figure 2



The word2vec similarity of the noun-verb pairing produced in each trial.

Figure 3

Lure strength: for each noun-verb production, each dot represents the cloze probability (in the current experiment) of the same verb in trials where the noun had the alternative case marker, e.g., for *thief-nom arrest*, the figure shows the cloze probability of *thief-acc arrest*. Elevated values indicate strong lures.



## References

- [1] Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161–163.
- [2] Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.
- [3] Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, 17(1), 117–129.
- [4] Chow, W.-Y., Smith, C., Lau, E., & Phillips, C. (2016). A “bag-of-arguments” mechanism for initial verb predictions. *Language, Cognition and Neuroscience*, 31(5), 577–596.
- [5] Chow, W.-Y., Lau, E., Wang, S., & Phillips, C. (2018). Wait a second! Delayed impact of argument roles on on-line verb prediction. *Language, Cognition and Neuroscience*, 33(7), 803–828.
- [6] Momma, S. M. (2016). *Parsing, Generation and Grammar*.
- [7] Chow, W.Y. & Kurenkov, I., Buffinton, J., Kraut, B., & Phillips, C. (2015). How predictions change over time: evidence from an online cloze paradigm. Poster presented at the 28th annual CUNY Human Sentence Processing Conference, Los Angeles, CA.
- [8] Matsuno, S., Mizuki, S., & Sakaki, T. (2019). 日本語大規模 SNS+Web コーパスによる単語分散表現のモデル構築 [Construction of a distributed word representation model using large-scale SNS + web-based Japanese corpora]. 人工知能学会全国大会論文集 [Proceedings of the Annual Conference of the Japanese Society of Artificial Intelligence], JSAI2019, 4Rin113-4Rin113.
- [9] Kuperberg, G. R. (2016). Separate streams or probabilistic inference? What the N400 can tell us about the comprehension of events. *Language, Cognition and Neuroscience*, 31(5), 602–616.