

What do eye movements in the visual world reflect? A case study from adjectives

Ciyang Qing, Daniel Lassiter, Judith Degen (Stanford University)

qciyang@stanford.edu

A common dependent measure used in visual-world eye-tracking experiments is the proportion of looks to a visually depicted object in a certain time window after the onset of the critical stimulus [1]. When interpreting such data, a common assumption is that looks to the object reflects the listener's belief that the object is the intended target referent. While this is intuitively plausible (at least for paradigms in which the task requires selecting a referent [2]), relatively little is known about how exactly the proportion of looks to an object is related to a listener's current belief about that object. Here, we test a simple, explicit linking hypothesis: ***the proportion of looks to an object reflects the probability that the listener assigns to the object of being the target***. To test this hypothesis, we supplement the eye-tracking data from [3] with an offline *incremental decision* task to measure participants' beliefs about the intended referent at various points in the unfolding sentence, and assess to what extent these beliefs predict the eye-tracking data.

Eye-tracking stimuli and data We used the eye movement dataset collected by [3], tested for online processing differences between maximum standard (e.g., *empty*) and relative (e.g., *big*) adjectives. Each scene consisted of 4 objects (see Fig. 1). Participants heard auditory stimuli like "Click on the empty cube". The target object exhibited a high degree of the property denoted by the adjective (e.g., emptiness) and for half of the scenes there was a contrast object exhibiting a low degree. In addition, there was a competitor object of a different kind that exhibited an even higher degree than the target. Finally, there were one or two distractors in the scene, depending on whether there was a contrast object. We reanalyzed data from three windows: prior, adjective, and noun (see Fig. 2).

Offline incremental decision task We recruited 100 participants via AWS Mechanical Turk. They were presented with the same scenes described above and were told that they were playing a game with another Turker, who sent a message to instruct them to click on one of the objects. They saw the critical sentence "Please click on the [adj] [noun]" incrementally and were asked to click on the object they thought was the intended referent after they saw the part of the sentence up to and including (i) the article "the," (ii) the adjective, and (iii) the head noun.

Results Fig. 3 shows the proportion of looks to each object against the proportion of clicks the object received in the incremental decision task. There was no significant correlation between click and eye data in the prior window for either maximum ($r = .029$, $p = .71$) or relative ($r = .003$, $p = .95$) adjectives. While there was a medium correlation between click and eye data for relative adjectives ($r = .37$, $p < .001$), there was only a weak correlation for maximum adjectives ($r = .17$, $p = .03 < .05$). In the noun windows, there was a strong correlation between click and eye data for both maximum ($r = .82$, $p < .001$) and relative ($r = .80$, $p < .001$) adjectives.

Discussion These results suggest that the degree to which an object is believed to be the referent is one factor that affects eye movements in referential tasks. In the talk we will discuss methodological implications of these results for experimental linguistics. We will also discuss extensions of the linking assumption to incorporate utterance surprisal and other factors that might influence eye movements.

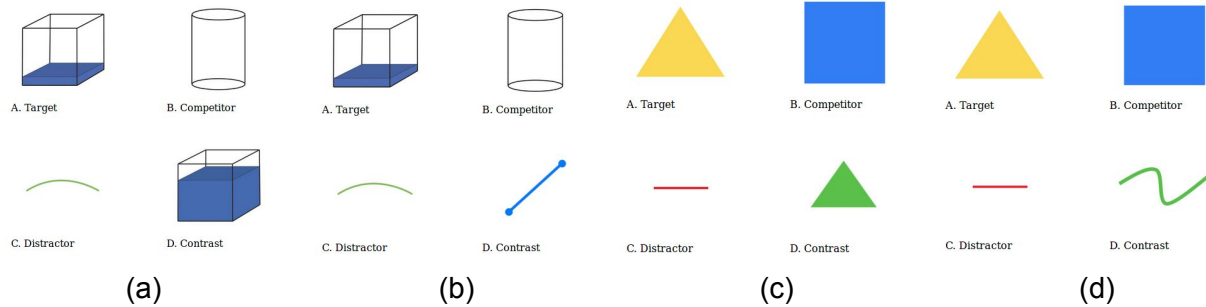


Fig.1 Visual stimuli from [3]: (1a) maximum adjective (*empty*), contrast, (1b) maximum adjective (*empty*), no contrast, (1c) relative adjective (*big*), contrast, (1d) relative adjective (*big*), no contrast.

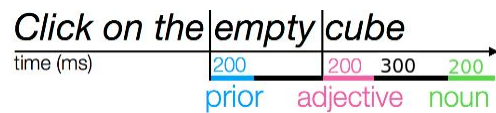


Fig. 2 Eye-tracking analysis time windows (200ms each).

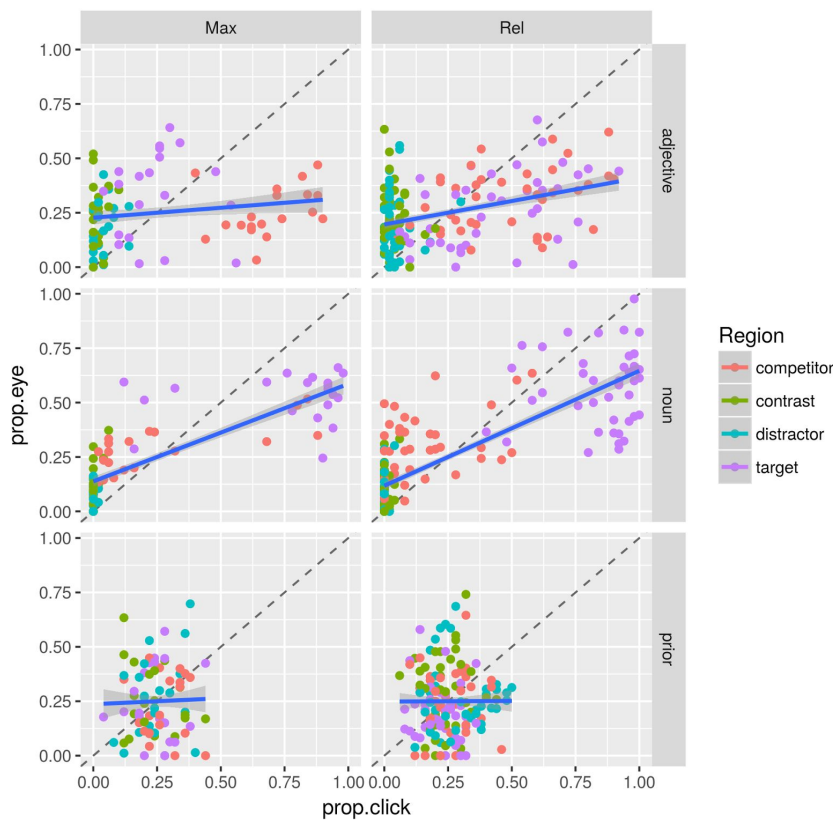


Fig. 3 Proportion of looks from [3] for maximum and relative adjectives in three different time windows (prior, adjective, noun) against click data from our offline incremental decision task.

References

- [1] Tanenhaus, M., Spivey, M., Eberhard, K., & Sedivy, J. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634.
- [2] Salverda, A. P., & Tanenhaus, M. (2017). The visual world paradigm. In A. de Groot & P. Hagoort (Ed.), *Research methods in psychology*.
- [3] Leffel, T., Xiang, M. & Kennedy, C. (2016). Imprecision is pragmatic: Evidence from referential processing. *SALT* 26, 836-854.

Acknowledgments

We are grateful to Tim Leffel, Ming Xiang, and Chris Kennedy for sharing their data with us.