How social contexts shape information gathering, attention, and memory during familiar language comprehension and novel word learning

Kyle MacDonald[*,1]

Abstract

Language comprehension in grounded contexts involves integrating visual and linguistic information through decisions about visual fixation. But when the visual signal also contains information about the language source – as in the case of written text or sign language – how do we decide where to look? Here, we hypothesize that eye movements during language comprehension represent an adaptive response. Using two case studies, we show that, compared to English-learners, young signers delayed their gaze shifts away from a language source, were more accurate with these shifts, and produced a smaller proportion of nonlanguage-driven shifts (E1). Next, we present a well-controlled, confirmatory experiment, showing that English-speaking adults produced fewer nonlanguage-driven shifts when processing printed text compared to spoken language (E2). Together, these data suggest that people adapt to the value of seeking different information in order to increase the chance of rapid and accurate language understanding.

Overview of the proposed work

The study of eye movements during language comprehension has provided fundamental insights into the interaction between conceptual representations of the world and the incoming linguistic signal. For example, research shows that adults and children will rapidly shift visual attention upon hearing the name of an object in the visual scene, with a high proportion of shifts occurring prior to the offset of the word (Allopenna, Magnuson, and Tanenhaus 1998; Tanenhaus et al. 1995). Moreover, researchers have found that conceptual representations activated by fixations to the visual world can modulate subsequent eye movements during language processing (Altmann and Kamide 2007).

The majority of this work has used eye movements as a measure of the output of the underlying language comprehension process, often using linguistic stimuli that come from a disembodied voice. But in real world contexts, people also gather information about the linguistic signal by fixating on the language source. Consider a speaker who asks you to "Pass the salt" but you are in a noisy room, making it difficult to understand the request. Here, comprehension can be facilitated by gathering information via (a) fixations to the nonlinguistic visual world (i.e., encoding the objects that are present in the scene) or (b) fixations to the speaker (i.e., reading lips or perhaps the direction of gaze).

But, this situation creates a tradeoff where the listener must decide what kind of information to gather and at what time. How do we decide where to look? We propose that people modulate their eye movements during language comprehension in response to tradeoffs in the value of gathering different kinds of information. We test this adaptive tradeoff account using two case studies that manipulate the value of different fixation locations for language understanding: a) a comparison of processing sign vs. spoken language in children (E1), and b) a comparison of processing printed text vs. spoken language in adults (E2). Our key prediction is that competition for visual attention will make gaze shifts away from the language source less valuable than fixating the source of the linguistic signal, leading people to generate fewer exploratory, nonlanguage-driven eye movements.

---

[*]Corresponding Author
   Email address: `kylem4@stanford.edu` (Kyle MacDonald)

Background literature

Completed work

Social cues modulate attention and memory in cross-situational word learning

Eye movements in American Sign Language comprehension reflect a tradeoff between gathering information about the linguistic signal and the nonlinguistic visual world

Children learning spoken language leverage visual information from social partners to facilitate language comprehension

Proposed work

Degraded linguistic signal and visual information

Eye movements for information seeking during novel word learning

Active selection of referents during social cross-situational word learning

Conclusion

References

Allopenna, Paul D, James S Magnuson, and Michael K Tanenhaus. 1998. "Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models." *Journal of Memory and Language* 38 (4). Elsevier: 419–39.

Altmann, Gerry, and Yuki Kamide. 2007. "The Real-Time Mediation of Visual Attention by Language and World Knowledge: Linking Anticipatory (and Other) Eye Movements to Linguistic Processing." *Journal of Memory and Language* 57 (4). Elsevier: 502–18.

Tanenhaus, Michael K, Michael J Spivey-Knowlton, Kathleen M Eberhard, and Julie C Sedivy. 1995. "Integration of Visual and Linguistic Information in Spoken Language Comprehension." *Science* 268 (5217). The American Association for the Advancement of Science: 1632.