Do eye movements go with fictive motion?

Teenie Matlock (tmatlock@psych.stanford.edu)
Department of Psychology, Stanford University
Stanford, CA 94305 USA

Daniel C. Richardson (richardson@psych.stanford.edu)
Department of Psychology, Stanford University
Stanford, CA 94305 USA

Abstract

Cognitive scientists interested in the link between language and visual experience have shown that linguistic input influences eye movements. Research in this area, however, tends to focus on literal language alone. In the current study, we investigate whether *figurative* language influences eye movements. In our experiment, participants viewed two-dimensional depictions of static spatial scenes while they heard either fictive motion sentences, such as *The palm trees run along the highway,* or non-fictive motion sentences, such as *The palm trees are next to the highway*. Overall, sentence type influenced participants' eye movements. Specifically, gaze duration on the figure (e.g., palm trees) was longer with fictive motion sentences than with non-fictive motion sentences. Our results demonstrate that figurative language influences visual experience. They provide further evidence that fictive motion processing includes mentally simulated motion.

Introduction

Imagine that you and a friend are sitting in a courtyard chatting. During the course of the conversation, you occasionally glance over at a long, thin stationary object on the ground. You assume the object is a tree branch or a walking stick until your friend says, "Oh! Look what slithered onto the courtyard." At that point, your perceptions and conceptions of the object dramatically change. The object goes from a piece of wood to a snake.

Situations like these—in which language influences the interpretation of objects and actions—are ubiquitous. The question addressed here is whether this influence is limited to *literal* language, or whether it also includes *figurative* language. We are especially interested in whether sentences such as *The road goes through the desert* or *The fence follows the coastline* (figurative because they include a motion verb but express no motion) affect eye movements. Our results suggest they do.

What We Know about Fictive Motion

Everyday language is replete with sentences such as (1a) and (1b). These are literal descriptions of static scenes.

(1a) *The road is in the desert*
(1b) *The fence is on the coastline*

Language is also full of sentences such as (2a) and (2b).

(2a) *The road goes through the desert*
(2b) *The fence follows the coastline*

These sentences are figurative because they contain a motion verb (e.g., *goes*, *follows*) but express no actual motion (Matlock, 2001).[1] They contrast with literal sentences with motion verbs, such as *The bus goes through the desert*, or *The herd of sheep follows the coastline*, which feature mobile agents that move from one point in space and time to another (Talmy, 1975; Miller & Johnson-Laird, 1976).

Despite the absence of actual movement with sentences such as (2a) and (2b), they have been claimed to involve *fictive motion*, an implicit mental simulation of "movement" through a construed scene (Talmy, 1983, 1996, 2000). On this view, the conceptualizer subjectively "scans" from one part of the scene to another, most notably, along the figure (i.e., prominent entity, subject noun phrase referent). For (2a), this means "moving" along the road, and for (2b), it means "moving" along the fence. According to the argument, fictive motion is a way to impose motion on what is otherwise a static scene. It enables the language user to compute information about the layout of a scene, for instance, a road in a desert in (2a), or a fence aligned with a coastline in (2b). Importantly, Talmy (1996) and other cognitive linguists do *not* maintain that fictive motion involves vivid imagery whereby the conceptualizer "sees" himself or herself (or any other animate entity) moving point by point along the figure in the scene being described. Instead, they take the motion to be relatively fleeting and tacit. (See also Langacker's *abstract motion*, 1986, 2000, and Matsumoto's *subjective motion*, 1996).[2]

At first, the claim that people simulate motion while processing descriptions of static scenes seems bizarre. Why would motion be processed, for instance, with sentences such as (2a) and (2b) when neither the road nor the fence is

---

[1] Like Rumelhart (1979) and Gibbs (1994), we do not maintain a hard and fast distinction between "literal" and "figurative". We simply use these terms here to operationalize two types of motion verb constructions: those that express motion and those that do not.

[2] Our study looks at just one type of fictive motion, Talmy's (2000) *co-extension path* fictive motion. There are many others.

capable of movement? Perhaps it is more reasonable to assume that such sentences are yoked to a purely static representation, as proposed by Jackendoff (2002). On this view, the representation underlying sentences such as (2a) and (2b) is static and atemporal. It is not unlike the representation underlying the literal sentences shown in (1a) and (1b), in which all points along the figure are activated simultaneously rather than incrementally.

Recent experimental work suggests that mental simulation, a fundamental part of cognition (e.g., Schwartz & Black, 1999; Freyd, 1983; Barsalou, 1999) generalizes to fictive motion. In several reading studies, Matlock (in press) investigated whether thinking about motion would affect fictive motion processing. In one study, participants read vignettes about fast or slow travel through a large-scale spatial region (e.g., driving in a desert), and then a fictive motion critical sentence, such as *The road goes through the desert*. Participants were quicker to read the critical sentences after reading about fast motion than they were after reading about slow motion. The same effects were also observed with easy versus difficult terrains, and with short versus long distances. Critically, the effect was *not* obtained with non-fictive motion test sentences at the end of the same stories, such as *The road is in the desert*. In sum, the results show that thinking about motion influences the processing of fictive motion sentences, but not the processing of comparable non-fictive motion sentences. They provide evidence that simulating motion is part of fictive motion understanding.

Matlock, Ramscar, and Boroditsky (2003a, 2003b) investigated whether engaging in thought about fictive motion would influence metaphoric construal of time in the way that engaging in thought about real motion has been shown to do (see Boroditsky, 2000; Boroditsky & Ramscar, 2002). In one experiment, participants were primed with fictive motion sentences or non-fictive motion sentences (e.g., *The tattoo runs along his spine* versus *The tattoo is next to his spine*) before reading this ambiguous question: "Next Wednesday's meeting has been moved forward two days. When is the meeting now that it has been rescheduled?"[3] When primed with fictive motion (congruent with an *ego-moving* construal), people were more likely to say, "Friday", suggesting they viewed themselves "moving" forward in time. When primed with non-fictive motion (congruent with a *time-moving* construal), people were more likely to say, "Monday", suggesting they viewed time as "moving" toward them. Another experiment issued this same question with one of two primes: *The road <u>goes</u> all the way to New York* (fictive motion away from conceptualizer), *The road <u>comes</u> all the way to New York* (fictive motion toward conceptualizer). The results indicated that people were more likely to respond "Friday" after the *away* prime and more likely to respond "Monday" with the *toward* prime. They suggest that people take a perspective and simulate motion when thinking about fictive motion, and that that in turn affects the way they perform abstract reasoning, such as reasoning about temporal movement.

Hence, people simulate motion when processing figurative sentences such as *The road runs along the coast*, and this naturally affects conceptual representation. Given this, we would like to know whether fictive motion also influences perceptual processing.

What Eye Movements Can Tell Us

Eye movements have been measured during a range of cognitive and perceptual activities (for review, see Richardson & Spivey, 2004). Scene perception has been studied in terms of the "bottom up" statistical properties of the image that attracts eye fixations, and in terms of the "top down" knowledge, beliefs or expertise that might affect how one person inspects a scene differently from another (for review, see Henderson, 2003). In a separate research tradition, eye tracking has been used to investigate reading, which engages both linguistic and perceptual processing (Rayner, 1998; Tinker, 1946).

Until recently the intersection between language and visual perception—looking at a scene and listening to a voice—had not been studied. The advent of head-mounted and remote eye tracking devices has allowed researchers to place participants in relatively rich, natural visual contexts and record how the eyes respond to spoken instructions and descriptions. Such experiments have yielded a surprisingly close integration between incremental linguistic processing and visual perception, demonstrating that eye movements to possible referents in the world are used to resolve temporary ambiguities in word recognition and syntactic structure (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), and to predict upcoming agents based on thematic role information (Altmann & Kamide, 2004).

Language has also been shown to modulate eye movements even when there is nothing to look at. In studies by Spivey and colleagues (Spivey & Geng, 2001; Spivey, Tyler, Richardson, & Young, 2000), people stared at a blank screen or closed their eyes and listened to a story that was spatially extended along an axis, for example, a story about a train going past, or a sequence of activities occurring on successive floors of a tall apartment block. While listening, participants' saccades tended to be extended along the horizontal or vertical axes that were consistent with those communicated in the story.

Whether it is in the presence of temporary ambiguity, or in the absence of visual input, linguistic input has been shown to influence eye movements. However, with the current surge of interest in language and vision (e.g., Henderson & Ferreira, 2004) one question has been left behind. What about figurative language? Does it influence eye movements? If so, how? This is an important question, for figurative language is not restricted to poetic or literary works. It is at least as pervasive in every day talk as literal

[3] The question is ambiguous because people are just as apt to answer Friday as they are Monday when the question is posed without any prime. (See Boroditsky, 2000; Boroditsky & Ramscar, 2002 for discussion.)

language, if not more pervasive (see Gibbs, 1994; Katz, Cacciari, Gibbs, & Turner, 1998; Lakoff, 1987).

Experiment

In the current study, participants viewed static depictions of scenes while they heard fictive motion and non-fictive motion sentences. Of interest was whether there would be differences between the eye movements that accompanied fictive motion sentences, such as *The palm trees run along the highway,* and those that accompanied non-fictive motion sentences, such as *The palm trees are next to the highway*. On the surface, the sentences convey similar information: Both include a linearly extended subject noun phrase reference (e.g., *palm trees*) and both describe a static spatial scene. However, the former has been argued to involve mentally simulated motion or scanning along the figure, but the latter has not. Would participants spend more time inspecting figures with fictive motion sentences than figures with non-fictive motion sentences? Longer gaze durations on regions of interest with fictive motion sentences would suggest mentally simulated motion or scanning.

Method

Participants

A total of 24 Stanford University psychology students participated for course credit. All had normal or corrected-to-normal vision.

Design

Gaze durations were recorded along the axis referred to by the subject of the sentence (the compatible region) and along the axis not referred to by the subject of the sentence (the incompatible region). Half the sentences included fictive motion language and half did not. Therefore, the experiment was a 2 x 2 design, with compatibility as one factor and sentence type as the other.

Stimuli

Sixteen pictures served as the primary visual stimuli. Each depicted a simple spatial scene and featured both a horizontally extended figure and a vertically extended figure, for example, a river extending from top to bottom, and a fence extending left to right. A further 16 pictures were used as filler items. All pictures were matched on level of color luminance.

Sixteen blocks of recorded English sentences served as primary stimuli. Each block contained two sentence pairs. Each pair included a FM-sentence (fictive motion sentence) and a comparable NFM-sentence (non-fictive motion sentence), for example, *The cord runs along the wall*, and *The cord is on the wall*. One sentence pair referred to the vertical object in a picture and the other referred to the horizontal object in that same picture. Figure 1 displays an example picture and its block of sentences. Sixteen sentences that described the filler pictures were also recorded.

We conducted three norming studies on our sentences and pictures. In the first, 57 Stanford undergraduates judged all FM- and NFM-sentences on a scale of 1 to 7, in which 1 indicated "makes no sense at all" and 7 indicated "makes good sense". The mean for all FM-sentences was 5.85 and the mean for all NFM-sentences was 6.02. A t-test showed no reliable difference between the two, $t(31) = 1.16$, $p > .1$. In the second norming study, 28 undergraduates rated pairs of FM- and NFM-sentences on how similar they were in meaning. They used a scale of 1 to 7, in which 1 indicated "not at all similar" and 7 indicated "very similar". The mean for all sentence pairs across all subjects was 6.04, with the highest average at 7 and the lowest average at 5.25. In a third norming study, 12 undergraduates judged our pictures and sentences on how well they went together. Overall, the sentence-picture combinations were judged as well-matched. The means were 6.63 FM-horizontal, 6.58 FM-horizontal, 6.53 NFM-horizontal, and 6.34 NFM-vertical. A one-way ANOVA yielded no difference, $F(3, 63) = .04$, $p > .1$, showing that they were equally good descriptions.

Together, the norming studies indicate that (a) all FM- and NFM-sentences were equally sensible in meaning, (b) all FM- and NFM-sentences described comparable information, and (c) all FM- and NFM-sentences were equally good descriptions of the pictures.

Horizontal landmark
FM *The books run along the wall*
NFM *The books are on the wall*

Vertical landmark
FM *The cord runs along the wall*
NFM *The cord is on the wall*

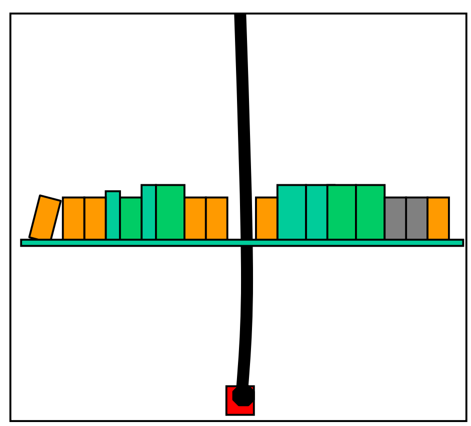

Figure 1. Example of picture with vertical and horizontal fictive motion and non-fictive motion sentences.

Apparatus
An ASL 504 remote eye tracking camera was positioned at the base of a 17" LCD stimulus display. Participants were unrestrained, and sat approximately 30" from the screen. The camera detected pupil and corneal reflection position from the right eye, and the eye tracking PC calculated point-of-gaze in terms of coordinates on the stimulus display. This information was passed every 33ms to a PowerMac G4 which controlled the stimulus presentation and collected gaze duration data. Prior to the experimental session, participants went through a 9 point calibration routine, which typically took between 2 and 5 minutes.

Procedure
Once a successful eye track was established, participants were told to "Look at the pictures and listen to the sentences." On each trial, a picture appeared 1000ms before the sentence began, and then remained in view for 2000ms after the sentence finished. There was a 2000ms inter-stimulus interval, during which participants saw a gray screen, roughly isoluminant with the pictures.

Following 4 practice trials, each participant was presented with a random sequence of 16 experimental and 16 filler trials. Each experimental picture was accompanied by one of four sentences that described the picture (e.g., vertical FM-sentence). Sentence presentation varied such that each participant heard 4 vertical FM-sentences, 4 horizontal FM-sentences, 4 vertical NFM-sentences, and 4 horizontal NFM-sentences.

Coding
The screen was partitioned into 17 non-overlapping regions of interest, corresponding to a central square, six squares spanning the horizontal axis, six squares spanning the vertical axis, and four squares in each corner (see Figure 2). During the period that the experimental picture was onscreen, total gaze durations in each region were recorded.

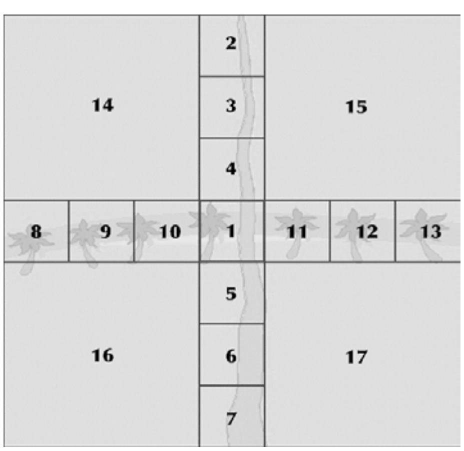

Figure 2. Grid defining relevant regions superimposed on grayed example picture.

Results and Discussion

For a quantitative analysis, we compared total gaze durations to the vertical region of the picture (regions 2 to 7) and the horizontal region of the picture (regions 8 to 13). Each accompanying sentence described either the horizontal or vertical element in the picture. Thus our data could be expressed in terms of gaze durations to the compatible and the incompatible regions. [4]

We conducted a 2 (compatibility) x 2 (sentence) ANOVA. There was a main effect of compatibility, indicating that our participants spent more time inspecting the compatible portion of the grid containing the figure described in the sentences, $F(1,23) = 5.51$, $p < .03$. There was also a main effect of sentence, revealing a reliable difference between inspection time for FM-sentences (M= 995) and inspection time for NFM-sentences (M=827), $F(1,23) = 10.18$, $p < .001$. Most importantly, there was a reliable interaction between these factors, $F(1,23) = 6.00$, $p < .03$, shown in Figure 3. Tukey's HSD revealed that the only cell that differed from all others was gaze duration to the compatible region with fictive motion sentences ($p < .01$).

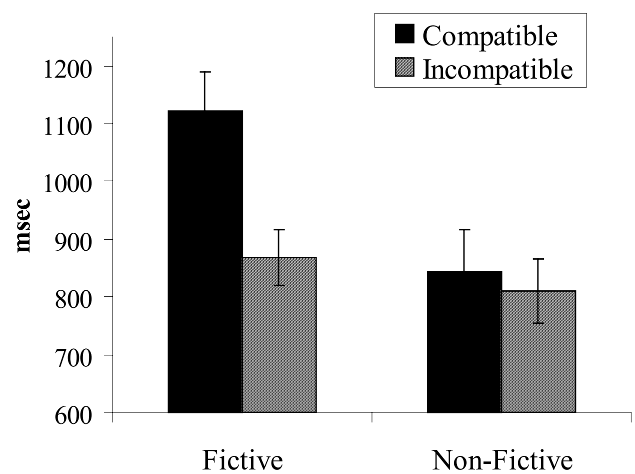

Figure 3. Gaze durations to compatible and incompatible picture regions only differed when the sentence employed fictive motion.

As predicted, people spent more time gazing at the region of a picture associated with the figure in fictive motion input than with the figure in non-fictive motion input, especially when the figure in the picture was compatible with the

[4] Example recordings of eye tracks can be seen at http://psychology.stanford.edu/~richardson/ficmot.

figure in the sentence. Taken together, our results show that fictive motion sentences had a consistent and dramatic effect on eye movements, most notably on the compatible region of interest.

General Discussion

Participants in our preliminary study spent more time inspecting the compatible region of interest in spatial scenes when they heard fictive motion sentences than when they heard non-fictive sentences. The results demonstrate that *figurative* language influences eye movements in consistent and predictable ways. The results are in line with other work on fictive motion (Matlock, in press; Matlock, Ramscar, & Boroditsky, 2003a, 2003b), and they suggest a dynamic mental representation that mirrors perception or enactment of motion (e.g., Barsalou, 1999, Glenberg, 1999).

One explanation for the results obtained here is that when our participants were presented with fictive motion input, they mentally simulated motion along the figure, and then their eye movements mirrored that internal simulation. For example, on hearing the sentence *The road runs through the desert*, participants conceptually "moved" along a road and then their eye movements enacted a congruent simulation. Another not incompatible explanation assigns a more active role to eye movements. It might be that participants' eye movements were central to simulation and building an appropriate representation of the figure. For example, on hearing the sentence *The road runs through the desert* and seeing a depiction of that scene, participants' eye movements allowed them to incrementally construct an appropriate model of the road. If this is the case, then perhaps eye movements allowed participants to simulate and compute some information about the scene externally (for related views, see Spivey, Richardson, & Fitneva, in press; Spivey, et al 2000).

Are there other explanations for longer gaze durations with fictive motion sentences? For instance, could it be that people activated the literal meaning of the motion verb in fictive motion sentences, and that that literal interpretation led to longer inspection times? Based on the results reported here, we cannot rule out this possibility entirely. But we would argue that our compatibility results suggest that this is not likely. Namely, if the verb alone – independent of the figurative meaning of the fictive motion sentence – brought on longer gaze durations, we would not have seen selective differences in the axis of orientation (vertical versus horizontal). After all, the motion verb alone provided no information about direction.

Our data show that figurative language, like literal language, influences eye movements. We argue that this is because fictive language evokes a dynamic mental simulation, and that this simulation determines how the visual system interprets and inspects the world. Further research will reveal how these simulations occur and the extent to which they mirror perception or enactment of physical motion in the world.

Acknowledgements

We thank Elsie Wang and Teresa-Bulygo for helping with our norming studies. We also thank Michael Spivey, Herbert Clark, Paul Maglio, Daniel Casasanto, and Gordon Bower for insightful comments.

References

Altmann, G. T. M., & Kamide, Y. (2004). Now you see it, now you don't: mediating the mapping between language and the visual world. In J. M. Henderson & F. Ferreira (Eds.), *Interfacing Language, Vision, and Action*. San Diego: Academic Press.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral & Brain Sciences, 22 , 577-660.*

Boroditsky, L. (2000). Metaphoric Structuring: Understanding time through spatial metaphors. *Cognition, 75*, 1-28.

Boroditsky, L. & Ramscar, M. (2002). The Roles of Body and Mind in Abstract Thought. *Psychological Science, 13(2)*, 185-188.

Freyd, J.J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics, 33*, 575-581.

Gibbs, R. W., Jr. (1994). *The Poetics of Mind : Figurative Thought, Language, and Understanding*. Cambridge, UK: Cambridge University Press.

Glenberg, A. M. (1999). Why mental models must be embodied. In G. Rickheit, & C. Habel (Eds.), *Mental models in discourse processing and reasoning*. New York, NY: North-Holland.

Henderson, J. M. (2003). Human gaze control in real-world scene perception. *Trends in Cognitive Sciences, 7*, 498-504.

Henderson, J. M., & Ferreira, F. (Eds.). (2004). *The integration of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Jackendoff, R (2002). *Foundations of language.* New York: Oxford University Press.

Katz, A.N., Cacciari, C., Gibbs, R.W., & Turner, M. (Eds.) (1998). *Figurative language and thought*. New York, NY: Oxford University Press.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago, IL: University of Chicago Press.

Langacker, R.W. (1986). Abstract motion. *Proceedings of the Twelfth Annual Meeting of the Berkeley Linguistics Society,* 455–471.

Langacker, R. W. (2000). Virtual reality. *Studies in the Linguistic Sciences, 29*, 77-103

Matlock, T. (in press). Fictive motion as cognitive simulation. *Memory & Cognition*.

Matlock, T. (2001). Fictive motion is real motion. Presentation at Seventh International Cognitive Linguistics Conference (ICLC), Santa Barbara.

Matlock, T., Ramscar, M., & Boroditsky, L. (2003a). The experiential basis of meaning. *Proceedings of the 25th Annual Conference of the Cognitive Science Society.*

Matlock, T., Ramscar, M., & Boroditsky, L. (2003b). The experiential basis of motion language. *Proceedings of Language, Culture, and Cognition: An International Conference on Cognitive Linguistics*. Braga, Portugal.

Matsumoto, Y. (1996). Subjective motion and English and Japanese verbs. *Cognitive Linguistics, 7*, 183-226.

Miller, G. & Johnson-Laird, P. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124(3),* 372-422.

Richardson, D. C., & Spivey, M. J. (2004). Eye Movements. In Wenk. G. & Bowlin, G. (Eds.) *Encyclopedia of Biomaterials and Biomedical Engineering*. Marcel Dekker, Inc.

Rumelhart, D.E. (1979). Some problems with the notion of literal meaning. In A. Ortony (ed). *Metaphor and Thought. Cambridge University Press*.

Schwartz, D. L., & Black, T. (1999). Inferences through imagined actions: Knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 116–136.

Spivey, M.J., & Geng, J.J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research/Psychologische Forschung, 65*, 235-241.

Spivey, M.J., Richardson, D.C., & Fitneva, S.A. (in press). Thinking outside the brain: Spatial indices to visual and linguistics information. In J. Henderson & F. Ferreira (Eds.), *Interfacing Language, Vision, and Action*. San Diego, CA: Academic Press.

Spivey, M.J., Tyler, M.J., Richardson, D.C., & Young, E.E. (2000). Eye movements during comprehension of spoken scene descriptions. *The Proceedings of the Twenty-second Annual Cognitive Science Society Meeting*, 487-492.

Talmy, L. (1975). Semantics and syntax of motion. In J. P. Kimball (Ed.), *Syntax and Semantics*, Volume 4 (pp.181–238). New York, NY: Academic Press.

Talmy, L. (1983). How language structures space. In H. Pick, & L.P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225-282). New York: Plenum Press.

Talmy, L. (1996). Fictive motion in language and "ception". In P. Bloom, M.A. Peterson, L. Nadel, & M.F. Garrett (Eds.), *Language and space* (pp. 211-276). Cambridge, MA: MIT Press.

Talmy, L. (2000). *Toward a Cognitive Semantics, Volume I: Conceptual Structuring Systems*. Cambridge: MIT Press.

Tanenhaus, M. K., Spivey Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*(5217), 1632-1634.

Tinker, M. A. (1946). The study of eye movements in reading. *Psychological Bulletin*.