Running head: EYE MOVEMENTS REFLECT GRAMMAR

Spontaneous eye movements during passive spoken language comprehension reflect grammatical processing

Stephanie Huette[1], Bodo Winter[1], Teenie Matlock[1], David Ardell[2], Michael Spivey[1]

1 Cognitive and Information Sciences, University of California, Merced
2 Quantitative Systems Biology, University of California, Merced
5200 North Lake Road
Merced, CA 95343

Corresponding Author:
Stephanie Huette
Cognitive and Information Sciences
University of California, Merced
5200 North Lake Road
Merced, CA 95343

Phone: (209) 228-7742
Email: shuette@ucmerced.edu

Abstract

Language is tightly connected to sensory and motor systems. Recent research using eye-tracking typically relies on constrained visual contexts, viewing a small array of objects on a computer screen. Some critiques of embodiment ask if people simply match their simulations to the pictures being presented. This study compared the comprehension of verbs with two different grammatical forms: the past progressive form (e.g., was walking), which emphasizes the ongoing nature of actions, and the simple past (e.g., walked), which emphasizes the end-state of an action. The results showed that the distribution and timing of eye movements mirrors the underlying conceptual structure of this linguistic difference in the absence of any visual stimuli. Thus, eye movement data suggest that visual inputs are unnecessary to solicit perceptual simulations.

Keywords: Language, Eye Movements, Linguistic aspect, Embodiment

Introduction

The capacity to think about past, present or future events is a fundamental cognitive ability (Zacks & Tversky, 2001). Language taps into this capacity by directing how one thinks of a particular event (Givón, 1992). Grammatical aspect has the ability to specify fine-grained temporal differences that are implied. With the sentence *John was going to the store* (past progressive) the unfolding of an event is emphasized, whereas with the sentence *John went to the store* (simple past) the end state is emphasized. This distinction is supported by linguistic research on aspect (Comrie, 1976; Dowty, 1977; Langacker, 1982), as well as psychological research on aspect (Madden & Zwaan, 2003; Magliano & Schleich, 2000; Matlock, 2011). Together, the work to date indicates that grammatical aspect influences the way we think about events, but little is in fact known about the underlying mechanisms and processes.

Behavioral and electrophysiological evidence demonstrate that sentences with progressive aspect activate richer detailed event knowledge, for instance, details about the location and the participants in a scene (Ferretti, Kutas & McRae, 2007; Carreiras, Carriedo, Alonso & Fernández, 1997). The progressive's emphasis on the ongoing nature of events has been argued to draw attention to the motion of described actions (Anderson, Matlock, Fausey & Spivey, 2008) and to facilitate congruent motor movement (Bergen & Wheeler, 2010). These properties are also reflected in co-speech gestures, which are more extended in the context of progressive language (Duncan, 2002; Parrill, Bergen & Lichtenstein, 2011). Non-progressive forms, on the other hand, have been found to direct attention to the completion of an event and the static endpoint of a movement (Madden & Zwaan 2003; Magliano & Schleich, 2000). The distinction between these two forms has important real-world consequences for how people interpret actions and ultimately how it affects attitudes and perceptions, including voting preferences (Matlock & Fausey, 2010) and eye witness testimony (Matlock, Sparks, Matthews, Hunter & Huette, 2012).

Most cognitive work on grammatical aspect has used tasks that are incapable of precisely measuring how it influences the understanding of event in real time. By recording eye movements on a blank computer screen during passive listening to sentences, we were able extract a data stream of eye movements that were co-occurring with the comprehension process, unbiased by task strategies. Eye-movement data have offered crucial guidance for theories of language processing in various specific contexts, such as reading (Rayner, 1998; Spivey & Tanenhaus, 1998), integrating diagrams with text (Hegarty & Just, 1993), following spoken instructions (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), and engaging in directed mental imagery in the absence of visual cues (Altmann, 2004, Spivey & Geng, 2001). Eye-tracking is an unobtrusive measure that collects multiple data points per experimental trial (saccadic eye movements and fixations to locations on a screen), but experiments using this technology typically involve tasks that require explicit judgments on visual or linguistic stimuli. As such, they are not spontaneous and may involve task demands. Our experiment addresses this concern.

Previous studies have shown that language comprehension interfaces with motion processing (Bergen & Wheeler, 2010; Zwaan & Taylor, 2006; Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermüller, 2005) and that eye movements on a blank screen reflect the spatial content of verbally described scenes (Spivey & Geng, 2001). Because movement processes reflect language in content-specific ways, and more specifically, because eye movements reflect linguistic content, we predicted the oculomotor system would interact with grammatical aspect paralleling the emphases on motion in each particular context. Based on the above-mentioned linguistic analyses of aspect and experimental studies, we predicted that eye movements should be noticeably shorter and more widely dispersed in past progressive. This prediction is motivated in the following way: If a series of past progressive sentences such as “He was going” induces focus on the ongoing properties of described events, this leads to more thoughts of motion which are intertwined with areas that drive eye movements. By contrast, a series of simple past sentences (“He went”) focus on the static, completed end-state of the described event resulting in longer fixation times as if staring at a static object or scene. In addition to the temporal signal, the spatial pattern would reveal a larger area of scanning in the past progressive.

Methods

Participants

Sixty-three right-handed, native English-speaking participants’ fixations and responses were monitored and recorded with an Eyelink II eye-tracking system in accordance with IRB standards. Participants received extra credit for participation in a social sciences course at University of California, Merced. All participants had normal or corrected-to-normal vision and reported having no hearing problems or language deficiencies.

Materials

Stories varied only by grammatical aspect and consisted of three to four sentences each. For example “John was on a bike ride yesterday. After he **sped / was speeding** across the valley, he **climbed / was climbing** a mountain range. Then he **pedaled / was pedaling** along a river and finally, he **coasted / was coasting** into a campground.” There was a slight difference between the total duration of the past progressive stories (186.9s) and the simple past stories (174.4s). To account for the possibility that more fixations in the progressive condition were due to this durational difference, we added 2 seconds of silence at the end of each sentence or clause boundary. This provides an equal time-window for the analysis across the total duration of each item. The sound start and end events were recorded along the eye movement data to allow parsing time into three windows: total time, linguistic stimulus only, and period of silence only. While there are other ways to control for time differences, this method has the advantage of allowing us to determine whether any effects of grammatical aspect persist after sentence completion.

Procedure

Eye movement data was recorded at 500 Hz. Participants completed a picture-viewing task that was unrelated to the main experiment. After completing that task, participants were informed they would next complete a task that would help them forget the pictures they had just viewed. Before the task, participants were told to keep the eyes open and look at the screen so recalibration is not necessary. Participants then listened to 24 short vignettes that in either the past progressive or the simple past condition. A total of 31 of the participants were randomly assigned to the progressive condition, and 32 of the participants to the simple past condition. There was no task while listening over the headphones, and a blank white screen was in front of them. After the end of the experiment, participants were asked what they believed the nature of the task was. No one reported having a hypothesis that grammar was the manipulation, or that they predicted a magnitude difference in eye movements. Most naïve hypotheses about the nature of the experiment included the first viewing task that was not a part of the experiment.

Results

Unless otherwise noted, reported results refer to the full period of each item (sentence period + the following silence). Participants showed differential spatial distributions of their eye movements as a function of grammatical aspect. In the non-progressive aspect condition (with simple past tense sentences), participants tended to fixate their eyes on the central portion of the blank screen throughout the experiment, with few looks to the periphery; see Fig. 1, upper row. By contrast, in the progressive aspect condition (with past progressive sentences), participants moved their eyes around in a wider area; Fig. 1, lower row. To standardize the comparison of eye-movement dispersion across these two conditions, each participant's fixation data were individually z-scored so that means were aligned and distributional characteristics were not an artifact of averaging variant means. Subsequently, real-time fixation data were pooled into cumulative distributions for progressive and non-progressive conditions. When the average time spent fixating each x,y pixel (i.e., dwell time) in the past progressive condition is subtracted from the average dwell times for every x,y pixel in the simple past condition, the differences are found in the center of the distribution around the mean. This results from a substantial difference in the kurtosis ("peakedness") of the two distributions. Cumulative kurtosis measures were higher in the simple past condition (x-axis: 11.2, y-axis: 11.8) than in the past progressive condition (x: 8.4, y: 7.3). Moreover, a two-sample Kolmogorov-Smirnov test along x and y screen dimensions revealed that non-progressive and progressive distributions differed significantly from one another along the x-axis (D=0.02056, df=62, p<.0001) and the y-axis (D=0.0599, df=62, p<.0001).

More detailed distributional analyses showed that the non-progressive simple past condition was associated with a greater proportion of sentences during which a participant fixated in only one location (using Monte Carlo simulated p-value, χ^2=56.1574, df=1, p<.0001), and also a greater proportion of sentences during which a participant fixated in no more than two locations (using Monte Carlo simulated p-value, χ^2=75.1473, df=1, p<.0001). In the progressive condition, participants swept out more across the visual plane as measured by Area of the Convex Hull (ACH) of standardized eye movements (truncating outliers with standardized ACH > 30 or ACH = 0, comparing medians by condition, Wilcox test W = 1949210, p = 8.647×10^{-12}), and moved their eyes for greater total distances as measured by Total Path Length (TPL) of standardized eye movements (excluding TPL = 0; comparing medians by condition, Wilcox test W = 2265549, p = 3.002×10^{-10}). All of these measures suggest that the eyes covered a wider area when listening to progressive sentences.

Not only did participants move their eyes around in a wider dispersion in the progressive aspect condition, they also produced briefer fixations in order to achieve that broad distribution. In the past progressive condition, fixation durations averaged 473 ms, whereas in the simple past (non-progressive) condition, fixation durations averaged 645 ms (independent samples t-test: t(61)=2.8, p=.006). Compared to other studies on grammatical processing or eye movements during language comprehension this is a quite large difference in fixation times. Thus, something as seemingly automated as how long the eyes remain stable in between saccadic eye movements is substantially influenced by the temporal emphasis implied by the grammar. This difference is present during the time segments in which speech is being played (past progressive mean: 543ms; simple past mean: 802ms; t(61)=3, p=.004). Importantly, this difference also persists when analyzing only the two-second silences in between each sentence (past progressive mean: 360ms; simple past mean: 428ms; (t(61)=2.7, ; p=.008). See Fig. 2. Note that the mean difference in fixation duration between the two conditions was much larger while the speech was playing (simple past: 802ms, past progressive: 544ms) than during the period of silence (simple past: 428ms, past progressive: 361ms). This speaks to the importance of the linguistic information in triggering oculomotor processes, as it suggests that eye movements may be more affected by grammar when they are co-occurring with these grammatical properties.

Discussion

As our results demonstrate here, there is nothing passive about passive listening: the eyes are actively moving in a way that reflects subtle grammatical differences in the linguistic input. The actual eye movement patterns are in line with what was predicted based on linguistic analyses of aspect and previous experimental work: past progressive appears to emphasize the ongoing motion of described actions and the details of described events, such that sentences with past progressive induce eye movements that are more widely dispersed – even while viewing a completely blank screen. These results suggest a smooth cascading of information from language processes in the brain all the way to oculomotor processes (Tanenhaus et al., 1995).

The results are consistent with theoretical accounts of real-time language processing that emphasize the role of sensorimotor properties in linguistic content (Zwaan & Taylor, 2006; Barsalou, 2009; Meteyard, Bahrami, & Vigliocco, 2007). These results also begin to hint at the underlying mechanism of perceptual simulation: rather than constructing a mental model, this appears to be a rapid cascade of motor firings that in the past have been associated with viewing more motion. Because language arrives and leaves so quickly, this kind of perceptual simulation would allow for rapid comprehension of implied and related ideas, and a memory trace of this simulation would help build a discourse context. Thus, perceptual simulation is accessing previously learned perceptual-motor information, and need not be solicited by explicit, concurrent visual stimuli.

These findings are in line with previous research that show how described events with detailed spatiotemporal parameters involve sensorimotor systems of the brain (Hauk et al., 2004; Pulvermüller, 2005; Meteyard et al., 2007). As a natural extension of previous work embodiment in language processing, we demonstrated that grammar affects a whole suite of different measurements connected to eye movements in a situation that minimizes task demands and mirrors real-world passive listening circumstances. This provides compelling evidence in favor of the view that the neural circuitry devoted to language is tightly connected with perceptual and motor areas of the brain, and begins to build a framework for investigating the mechanisms of perceptual simulation.

References

Altmann, G. (2004). Language-mediated eye movements in the absence of a visual world: the 'blank screen paradigm'. *Cognition*, 93, B79-B87.

Anderson, S. E., Matlock, T., Fausey, C. M., & Spivey, M. J. (2008). On the path to understanding the on-line processing of grammatical aspect. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2253-2258). Austin, TX: Cognitive Science Society.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1281-1289.

Bergen, B., & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain & Language*, 112, 150-158.

Carreiras, M., Carriedo, N., Alonso, M. A., & Fernández, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during. *Memory & Cognition*, 25, 438-446.

Comrie, B. (1976). *Aspect*. Cambridge, MA: Cambridge University Press.

Dowty, D. R. (1977). Toward a semantic analysis of verb aspect and the English 'imperfective' progressive. *Linguistics and Philosophy,* 1, 45-77.

Duncan, S. D. (2002). Gesture, verb aspect, and the nature of iconic imagery in natural discourse. *Gesture*, 2, 183-206.

Ferretti, T. R., Kutas, M., & McRae, K. (2007). Verb aspect and the activation of event knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 182-196.

Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics,* 30, 5-55.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41, 301-307.

Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717-742.

Langacker, R. W. (1982). Remarks on English Aspect. In P. J. Hopper (Ed.), *Tense-Aspect: Between Semantics & Pragmatics* (pp. 265-304). Amsterdam: John Benjamins.

Madden, C. J., & Zwaan, R. A. (2003). How does verb aspect constrain event representations? *Memory & Cognition*, 31, 663-672.

Magliano, J. P., & Schleich, M. C. (2000). Verb aspect and situation models. *Discourse Processes*, 29, 83-112.

Matlock, T. (2011). The conceptual motivation of aspect. In K. Panther, & G. Radden (Eds.), *Motivation in Grammar and the Lexicon* (pp. 133-147). Amsterdam: John Benjamins.

Matlock, T., & Fausey, C. M. (2010). Can grammar win elections? *Political Psychology*, 32, 563-574.

Matlock, T., Sparks, D., Matthews, J. L., Hunter, J., & Huette, S. (in press). Smashing new results on aspectual framing: How people describe car accidents. *Studies in Language* (special issue).

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion words: Language affects low-level visual perception. *Psychological Science*, 18, 1007-1013.

Parrill, F., Bergen, B., & Lichtenstein, P. (2011). Grammatical aspect in language production: Using gesture to reveal event representations. In L. Carlson, C. Hoelscher, & T. F.

Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3223-3228). Austin, TX: Cognitive Science Society.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, 6, 576-582.

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

Spivey, M. J., & Geng, J. (2001). Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research*, 24, 235-241.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and* Cognition, 24, 1521-1543.

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, 1632-1634.

Taylor, L. J., & Zwaan, R. A. (2008). Motor resonance and linguistic focus. *Quarterly Journal of Experimental Psychology*, 61, 894-904.

Zacks, J.-M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin,* 127, 3-21.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. *Journal of Experimental Psychology: General*, 135, 1-11.

Acknowledgements

We thank Elaine Lai for assistance with data collection. We thank James Grama for recording our stimuli. This work was supported by NSF grants BCS-0721297 and BCS-1031903.

Figures

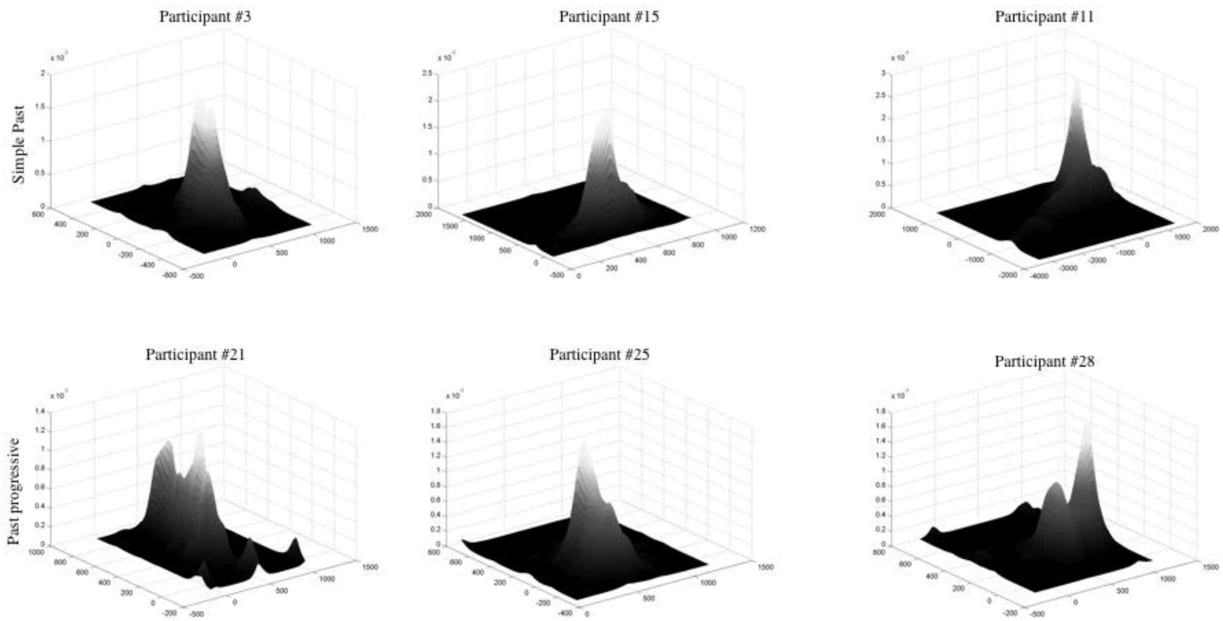

Fig. 1. Representative sample of individual fixation patterns, revealing a wider spread of eye movements in the past progressive condition as opposed to the simple past condition. The vertical axis shows total time spent fixating a given x,y location on the blank white screen. Each plot was z-scored. To more accurately represent dense areas, bivariate data was smoothed via a procedure appropriate for skewed data sets, for visual presentation only.

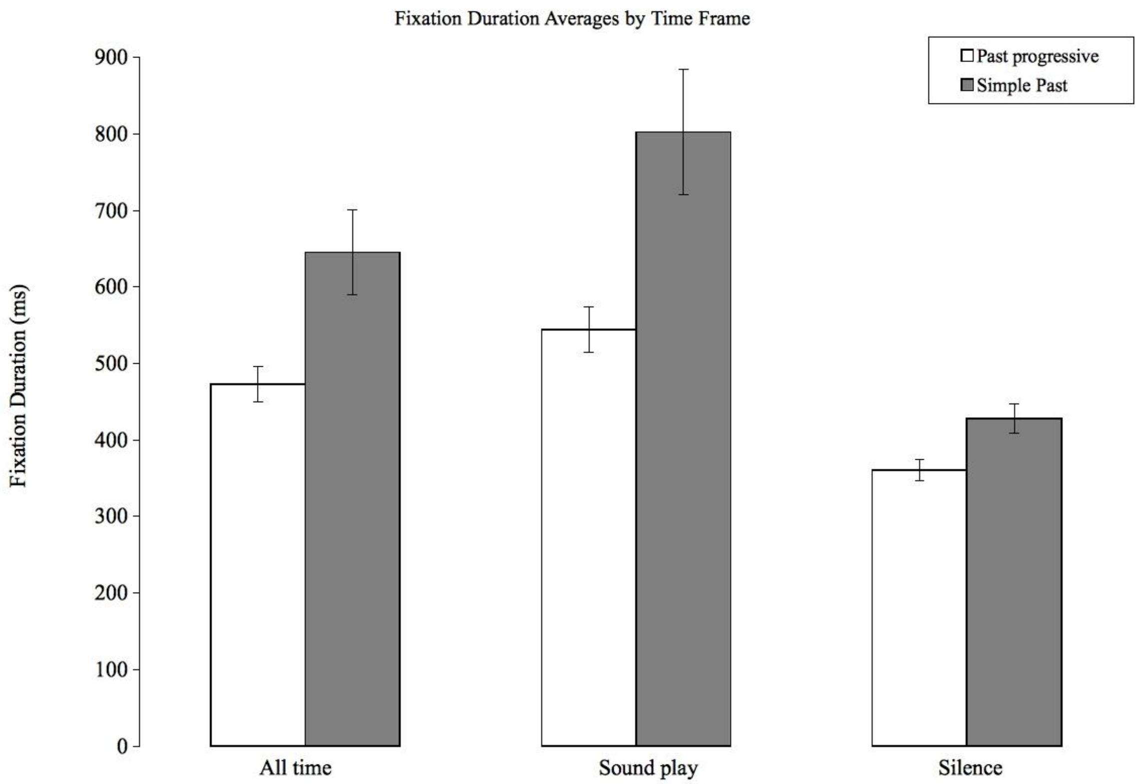

Fig. 2. Fixation duration averages in simple past and past progressive conditions by time frame. All time is both sound playing and silence data pooled. The average fixation duration is shorter in the progressive condition where the grammar implies an emphasis on motion. Error bars are SEM.