Motion in Vision and Language: Seeing Visual Motion can influence Processing of Motion Verbs

Carolin Dudschig (carolin.dudschig@uni-tuebingen.de)
Department of Psychology, Schleichstr. 4
72076 Tübingen, Germany

Jan Souman (jansouman@hotmail.com)
Department of Psychology, Schleichstr. 4
72076 Tübingen, Germany

Barbara Kaup (barbara.kaup@uni-tuebingen.de)
Department of Psychology, Schleichstr. 4
72076 Tübingen, Germany

Abstract

In contrast to symbolic models of language understanding, embodied models of language comprehension suggest that language is closely connected with visual and motor processing. In the current study we show that motion words, such as *rise* or *fall*, are processed faster if displayed against a background of compatible motion (e.g., upward vs. downward random dot motion with 60% motion coherence). However, this interaction between semantic processing and visual processing only occurred if the word and the motion display were presented simultaneously. If the visual motion display was short-lived and occurred 100 or 200 ms after word-onset, no interactions between language and visual motion were found. We suggest that only in situations that do not allow ignoring or strategically suppressing the visual motion display, supra-threshold visual motion can affect language comprehension.

Keywords: Language processing; motion verbs; vision; visual motion processing; embodiment; grounding.

Introduction

Embodied models of language understanding propose a close connection between language and perceptuomotor processes in the brain (e.g., Barsalou, 1999). Recently, compelling evidence supported the close association between language and other cognitive functions (e.g., Zwaan, Stanfield & Yaxley, 2002). In the motor domain converging evidence suggests that language facilitates compatible motor actions (e.g., Glenberg & Kaschak, 2002) and that language comprehension involves cortical motor areas that are also involved in performing the described actions (e.g., Hauk, Jonsrude, & Pulvermüller, 2004). For example, Glenberg and Kaschak showed that processing sentences such as "*Close the drawer*" can interfere with motoric responses incompatible with the motion implied in the sentence (e.g., arm movement towards my body). Similar effects have been reported in studies using motion verbs (e.g., *rise*, *climb*) or nouns implicitly implying a location (e.g., *bird* vs. *shoe*), whereby upward verbs and nouns facilitate upward arm movements (Dudschig, Lachmair, de la Vega, De Filippis, & Kaup, 2012a; Lachmair, Dudschig, De Filippis, de la Vega & Kaup, 2011). In contrast to the effects of language on motor processing, in the perceptual domain there is rather mixed evidence regarding the relation between language and visual processing. In particular, evidence regarding the influence of non-linguistic factors on language processing is rare. This direction of cause is particularly important, as these findings would suggest that mechanisms underlying non-linguistic processes are required and recruited during language processing.

Studies in the visual domain typically investigate the influence of language on perceptual detection or discrimination tasks. For example, it has been shown that words referring to entities with a typical location (e.g., *hat* vs. *shoe*) can influence visual target perception in upper or lower screen locations (e.g., Dudschig, Lachmair, de la Vega, De Filippis, & Kaup, 2012b; Estes, Verges & Barsalou, 2008). Similar results have been reported for valence words (e.g., Meier & Robinson, 2004) and religious concepts (e.g., Chasteen, Burdzy & Pratt, 2010). Additionally, there have been studies demonstrating that visual simulation can also occur during sentence processing and subsequently affect visual discrimination performance (Bergen, Lindsay, Matlock & Narayanan, 2007). Recently, it has been shown that not only visual discrimination performance but also eye-movements can be affected by words referring to entities in the upper or lower field of vision (Dudschig, Souman, Lachmair, de la Vega, & Kaup, 2013). More specifically, upward saccades are faster following words referring to entities in the upper visual field (e.g., *bird*) and in contrast, downward saccades are faster following words referring to entities in the lower visual field (e.g., *shoe*). Importantly, the relation between language and visual processing was also reported in the other causal direction: Perceiving visual motion patterns can affect language processing. For example, Kaschak, Madden, Therriault, Yaxley, Aveyard, Blanchard and Zwaan (2005) first reported the effects of visual motion perception on language comprehension. In their study, participants viewed visual motion patterns (e.g., upward vs. downward moving horizontal stripes on a screen) and at the same time had to listen to sentences and perform a sensibility judgment task. The results showed that reading times were slower when the visual motion (e.g., upwards pattern) matched the motion direction implied by the sentence (e.g., "*The rocket blasted off*"). The authors concluded that language processing demands access to visual processing resources. If these visual processing resources are engaged by the processing of motion patterns, sentence understanding can be impaired.

Interestingly, studies investigating the effect of visual motion percepts on single word comprehension reported opposing results. Meteyard, Zokaei, Bahrami and Vigliocco (2008) analyzed how the understanding of motion verbs (e.g., *rise* vs. *fall*) is influenced by activation of motion-responsive visual brain areas. In their study, motion verbs were presented on a screen together with a short-lived (200ms) visual motion pattern, whereby the visual motion pattern was noisy to a greater or lesser extent. In the near-threshold condition, the motion display was presented at a coherence level that made it difficult for the participants to detect the motion direction of the motion pattern. In the above-threshold condition, motion coherence was set to a level that clearly allowed classification of the motion direction (upward vs. downward moving pattern). Participants had to perform a lexical decision task. The results showed that near-threshold motion patterns facilitated processing of words implying compatible motions (e.g., *rise* was faster processed if presented together with a near-threshold upward motion). In the other experiments where the visual motion was set to above-threshold levels no effect of visual motion perception on language processing was observed. The authors suggested that visual motion activates motion-responsive areas in the brain (MT +). However, this activation can be suppressed by top-down control mechanisms in the case of above-threshold motion coherence only. Thus, only in near-threshold motion patterns the motion information resulted in interactions with semantic language processing. In contrast, in the case of above-threshold visual motion pattern top-down control was recruited and suppressed this visual activation. Importantly, in the study by Meteyard et al. the visual motion patterns were presented very briefly (200ms) in contrast to 35 sec visual motion percepts in the study of Kaschak et al. (2005). Taken together there is mixed evidence regarding the influence of visual motion perception on language processing. On the one side, above-threshold and long-lasting visual motion can influence sentences processing (Kaschak et al., 2005), on the other side, only near-threshold visual-motion patterns affected lexical access to single words (Meteyard et al., 2008). Thus, it remains open whether above-threshold visual motion can interact with semantic language processing on a word-level.

In the current study we investigate whether single-word processing can be affected by above-threshold visual motion if visual motion patterns are presented from word onset until response. Such findings would be important for the embodied model of language understanding, as they would suggest convergence in the empirical evidence in favor of the model, and suggest that both word and sentence processing are influenced similarly by co-occurring visual motion. In order to test this we adapted the visual motion displays used by Meteyard et al. (2008) and created above-threshold random dot motion displays, that clearly allowed classification of the motion as an upward or downward directed motion. Additionally, we manipulated the *stimulus onset asynchrony* (SOA) between the word display and the visual motion display. In the 0ms SOA condition the word and visual motion were displayed simultaneously. In the 100ms SOA condition first the word was displayed and after 100ms delay the visual motion pattern appeared. Similarly, in the 200ms SOA condition, the visual motion pattern followed the word display by 200ms. Importantly, only in the 0ms SOA condition word and motion display fully overlapped. Thus, in this condition the simultaneous presentation of word and motion display minimizes the possibility of the participants to ignore the visual motion display. We expected that in conditions were participants were constantly exposed to visual motion during the lexical decision task, visual motion will most strongly influence semantic language processing.

Method

Participants

Eighteen right-handed psychology students from the University of Tübingen took part in this experiment (M_{age} = 24.39, 16 female) for monetary reward or course credit.

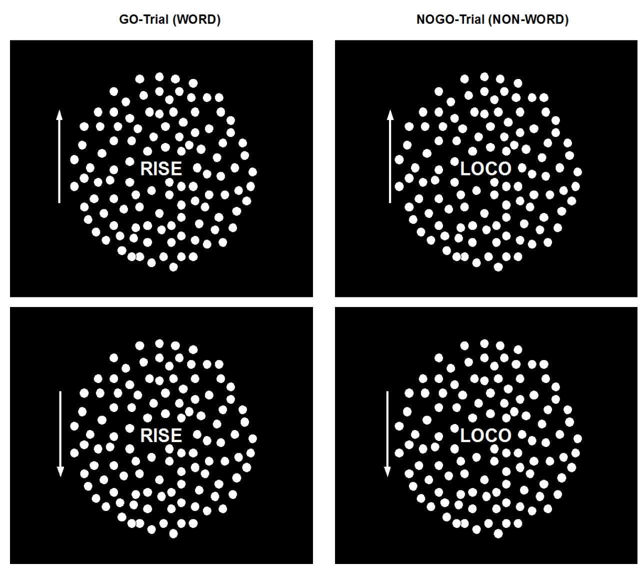

Figure 1: Trial examples for Go-Trials (word) and NoGo Trials (non-words). Visual motion was either compatible to the motion implied by the verb (top-left display) or incompatible (bottom-left display). Arrows illustrate visual motion direction and were not displayed in the actual experimental setup.

Stimuli & Apparatus

The experiment took place in a sound-attenuated booth. Participants viewed the screen from a 60cm viewing distance. Experimental procedure was implemented in MATLAB R2010a, Psychtoolbox, 3.0.8.

Words Twenty-four German verbs[1] referring to upwards motion and 24 verbs referring to a downwards motion were used as experimental stimuli. Upwards and downwards motion verbs did not differ in length (M_{up} = 8.74 (SD = 1.18), M_{down} = 8.35 (SD = 2.01), $t(44)$= .69, p = .49). Word frequency was retrieved from the Leipziger Wortschatzportal, upwards and downwards motion verbs did not differ in word frequency (M_{up} = 1886.17 (SD = 3545.31), M_{down} = 1667.70 (SD = 3134.64), $t(44)$ = .22, p = .83. Additionally, 48 pronounceable non-words were constructed. Therefore we used a different set of German verbs and permuted and exchanged various letters.

Visual Motion Patterns Visual motion patterns were adapted from Meteyard et al. (2008) with some adjustments, in order to make the motion clearly visible to the participants. 1000 moving dots were included in each display moving at a speed of 20°/s. Dot size was 0.1°. Dots were presented within an aperture of approximately 15cm diameter. Figure 1 shows examples of compatible and incompatible visual motion trials.

Procedure & Design

Each experimental trial started with the presentation of a fixation cross in the middle of the screen for the duration of 500ms (size: 20 pixels). Then, either a word or a non-word replaced the fixation cross. Words were presented in Arial font with a size of 0.5° x 2.5° visual angle. In the 0ms SOA condition, the visual motion pattern was presented together with the word. In the 100ms and 200ms SOA conditions, the visual motion pattern followed word onset by 100 or 200 ms, respectively. Words and visual motion were presented until response. Participants had to press the space bar if they decided that the displayed stimulus is a word and withhold response in case of non-word trials. If no response was recorded within 1500ms the next trial started automatically. The inter-trial-interval was 500ms. 20 Practice trials were conducted using a separate set of verbal stimuli. The experiment consisted of 576 Go-Trials (word trials) and 576 NoGo-Trials (non-word trials). Each of the 48 words was presented four times in each SOA condition (twice with an upward motion pattern and twice with a downward motion pattern). The experimental design was a within-subject design, with the factors SOA (0, 100, 200ms), visual motion (upward, downward) and word direction (upward, downward).

[1] Exemplary German verbs denoting to upwards motion: *steigen* to rise), *erhöhen* (to increase), *klettern* (to climb), *wachsen* (to grow), *hissen* (to hoist), *erheben* (to lift) etc.. Exemplary German verbs denoting to downwards motion: *fallen* (to fall), *sinken* (to sink), *tauchen* (to dive), *tropfen* (to drip), *landen* (to land), *schütten* (to pour), *einstürzen* (to collapse) etc.

Results

All NoGo-Trials and erroneous trials were excluded from analysis. Error exclusion reduced the dataset by 1.40 %. Additionally, outliers were excluded from reaction time (RT) analysis, with a criterion of 4 SD reducing the dataset by less than 0.43%. The lexical decision times were analyzed in two repeated measures ANOVAs. In the first ANOVA participant was the random-factor (F_1: by-participant analysis) and in an additional ANOVA the stimulus word served as random-factor (F_2: by-item analysis).

Reaction time results are displayed in Figure 2. There was a main effect of word direction in the by-participant analysis only, $F_1(1,17) = 13.06$, $MSE = 834$, $p < .01$, $F_2(1,46) = 1.56$, $MSE = 11520$, $p = .22$, with responses to downwards word (624 ms) being faster than to upwards words (639 ms). There was no effect of visual motion, $F_1(1,17) = 0.12$, $MSE = 652.8$, $p = .74$, $F_2(1,46) = 0.27$, $MSE = 731.4$, $p = .60$, nor of SOA, $F_1(2,34) = 0.59$, $MSE = 721.2$, $p = .56$, $F_2(2,92) = 2.53$, $MSE = 896.2$, $p = .09$. There was no interaction between visual motion and SOA, $F_1(2,34) = 0.54$, $MSE = 454.5$, $p = .59$, $F_2(2,92) = 0.64$, $MSE = 901.3$, $p = .53$. There was no interaction between word direction and SOA, $F_1(2,34) = 2.56$, $MSE = 486.4$, $p = .09$, $F_2(2,92) = 0.34$, $MSE = 896.2$, $p = .71$. There was no interaction between word direction and visual motion direction, $F_1(1, 17) = 3.15$, $MSE = 573.8$, $p = .09$, $F_2(1,46) = 0.62$, $MSE = 731.4$, $p = .44$. Importantly, there was a significant three-way-interaction between word direction, visual motion and SOA, $F_1(2, 34) = 3.56$, $MSE = 456.7$, $p < .05$, $F_2(2,92) = 3.03$, $MSE = 901.3$, $p = .05$. Separate analysis of the SOA conditions showed, that the three way interaction was due to the interaction between word direction and visual motion being significant for the 0ms SOA condition only, $F_1(1,17) = 8.64$, $MSE = 585$, $p < .01$, $F_2(1,46) = 6.43$, $MSE = 790$, $p < .05$ and not for the 100ms SOA, $F_1(1,17) = 0.01$, $MSE = 483.4$, $p = .94$, $F_2(1,46) = 0.96$, $MSE = 839.9$, $p < .033$ or the 200ms SOA, $F_1(1,17) = 0.00$, $MSE = 491.1$, $p = .97$, $F_2(1,46) = 0.04$, $MSE = 904.3$, $p < .84$. In summary, visual motion direction did interact with lexical processing. However, this was only in trials were word and visual motion display fully overlapped (0ms SOA condition). Post-hoc tests showed that this effect was due to faster classification of words referring to upward motion (e.g., *rise*, *climb*) if presented on the background of an upward motion in contrast to a downward motion, $t_1(17) = -2.27$, $p < .05$, $t_2(23) = -2.40$, $p < .05$. In contrast word referring to a downward motion (e.g., *fall*, *drip*) were faster classified if presented on the background of a downward motion, this was reflected in a trend in the by-subject analysis, $t_1(17) = 1.93$, $p = .07$, $t_2(23) = 1.24$, $p = .22$.

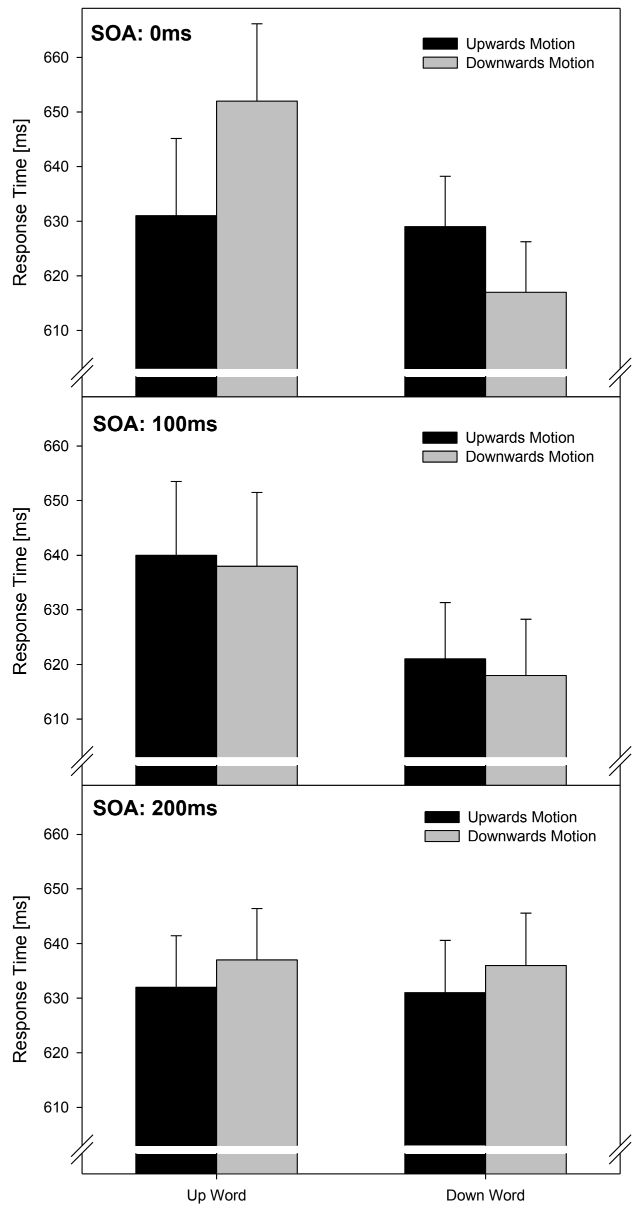

Figure 2: Reaction time results for the lexical decision task, separately for the three SOA conditions, the word direction and the visual motion direction. Error bars represent confidence intervals for within-subject designs according to Loftus and Masson (1994).

Discussion

Converging evidence suggests that language processing is closely related to other cognitive functions and can affect visual and motor processing. Interestingly, some studies also report an effect of motor processes (e.g., Glenberg, Sato, & Cattaneo, 2008) or visual processing (e.g., Kaschak et al., 2005; Meteyard et al., 2008) on language comprehension, suggesting direct involvement of visual and motor processes during language understanding. Kaschak et al. (2005) reported that visual motion perception (e.g., downwards motion) interferes with understanding of sentences that imply compatible motions (e.g., "*The confetti fell on the parade*"). In contrast, Meteyard et al. showed that near-threshold visual motion facilitates lexical access to words that imply compatible motion directions. In the current experiment, we addressed the question whether above-threshold visual motion can affect lexical processing of motion verbs if the participants have no possibility to strategically ignore the visual information. Indeed, our results showed that in conditions where word display and visual motion display occurred simultaneously (0ms SOA) and persisted throughout the trial, visual motion patterns did interact with lexical processing of the motion verbs. More specifically, we found that upward motion words (e.g., *rise*) are processed faster if displayed against the background of an upward motion than against downward motion, and the opposite holds for downward motion verbs.

To our knowledge, these findings are the first that show an effect of above-threshold visual motion on single-word processing. But why do we not find interference effects as reported by Kaschak et al. (2005)? First of all, single-word processing might differ regarding the mechanisms how visual processing resources are activated during reading, thus language-vision interactions might occur at different time-points or processing stages. Indeed, previous studies showed that timing can play a crucial role and may determine whether facilitation or interference effects are found (e.g., Boulengner, Roy, Paulignan, Deprez, Jeannerod & Nazir, 2006). Additionally, in our study we used motion patterns that were very different from Kaschak et al. (moving dot patterns vs. moving bars) and our moving dot patterns were only displayed during each trial. In contrast, Kaschak et al. (2005) displayed motion for as long as 35s and motion display extended between trials. Moreover, sentences were presented auditory in Kaschak et al.'s study. Thus, differences in task parameters and language material might results in facilitation effects in our study. Indeed, we adapted our visual motion patterns from the study of Meteyard et al. who also reported facilitation effects in case of single-word processing and lexical decision tasks.

This directly leads to the next question: Why do we find a facilitation effect despite using above-threshold motion patterns that can be clearly classified as upward or downward moving motion pattern? In the study of Meteyard et al. these influences of visual motion on language understanding were only observed for near-threshold motion patterns. Previously it has been suggested that the influence of task-irrelevant sub-threshold motion patterns on task performance is stronger than the influence of supra-threshold motion patterns (Tsushima, Sasaki, & Watanabe, 2006). The authors suggested that sub-threshold motion patterns are processed in the visual cortex similar to supra-threshold motion patterns; however in contrast to supra-threshold motion patterns sub-threshold motion patterns do not automatically result in recruitment of inhibitory control from the lateral prefrontal cortex (LFPC) in order to inhibit the visual cortex activation (in MT+) and thus reduce the influence of the motion percept on responding. The fact that we do find an influence of supra-threshold visual motion patterns on lexical decision task might have several implications. First, as our motion display occurred throughout the whole trial participants might fail to recruit sufficient top-down control mechanisms in order to fully suppress the influence of the visual motion on performance. Additionally, if language processing and visual processing are directly related, small activation in the visual cortex might also be sufficient to influence language processes. Thus, due to top-down control inhibitory control from the LPFC that suppressed visual motion activation, the effects in our study might be rather small. Additionally, in the 100ms and the 200ms SOA condition, the LPFC suppression mechanisms on the MT+ activation might be stronger, as it might be easier to suppress the influence of a visual motion display that is delayed in onset to the critical stimulus. Further studies will be required to fully understand the interplay between the language and the visual system and the critical time intervals during language processing, where this interaction occurs.

In summary, our findings have several implications. First, our results suggest that visual motion can also affect language processing if visual motion is presented above-threshold. Second, these findings pose a challenge to some findings in the motor domain. Typically, in motor tasks participants can see their arm or hand motion. Thus, if participants are instructed to perform a lexical decision task, decision times might be faster in compatible directions, because the visual input from the moving arm or hand will interact with lexical processing. Thus, given that our findings show that word processing interacts with visual motion perception, some findings in the motor domain might also be explained by perception of actual arm or hand movements. In future studies interactions between motor action and language need to be considered carefully, as potentially also being influenced by visual motion perception. In summary, our findings show that language processing and visual processing are closely interrelated. In paradigms, where participants cannot ignore or actively avoid motion perception, language processing can be facilitated by compatible visual motion.

Acknowledgments

We thank our student assistants for help with data collection. This project was supported by a Margarete-von-Wrangell Fellowship appointed to Carolin Dudschig (European Social Fund and the Ministry Of Science, Research and the Arts Baden-Württemberg) and by the SFB833/B4 project of Barbara Kaup (German Research Foundation).

References

Barsalou, L.W. (1999). Perceptions of Perceptual Symbols. *Behavioral and Brain Sciences, 22,* 637-660.

Bergen. B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*, 733-764.

Boulenger, V., Roy, A. C., Paulignan, Y., Deprez, V., Jeannerod, M., & Nazir, T. A. (2006). Cross-talk between language processes and overt motor behavior in the first 200 ms of processing. *Journal of Cognitive Neuroscience, 18*, 1607–1615.

Chasteen, A.L., Burdzy, D.C., & Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia*, *48*, 627-630.

Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). From top to bottom: spatial shifts of attention caused by linguistic stimuli. *Cognitive Processing*, *13*, 151-154.

Dudschig, C., Lachmair, M., de la Vega, I., De Filippis, M., & Kaup, B. (2012). Do task-irrelevant direction-associated motion verbs affect action planning? Evidence from a Stroop paradigm. *Memory & Cognition*, *40*, 1081-1094.

Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., & Kaup, B. (2013). Reading "Sun" and Looking Up: The Influence of Language on Saccadic Eye Movements in the Vertical Dimension. *PloS ONE*, *8*, e56872.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the object's typical location. *Psychological Science*, *19*, 93–97.

Glenberg, A.M. & Kaschak, M.P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558-565.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology*, *61*, 905-919.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 303-307.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*, B79-B89.

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: automatic activation of location information during word processing. *Psychonomic Bulletin & Review*, *18*, 1180-1188.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476-490.

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, *18*, R732-R733.

Tsushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. *Science*, *314*, 1786-1788.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*, 168-171.