Abstract language comprehension is incrementally modulated by non-referential spatial information: evidence from eye-tracking

Ernesto Guerra (ernesto.guerra@uni-bielefeld.de)
Pia Knoeferle (knoeferl@cit-ec.uni-bielefeld.de)
Cognitive Interaction Technology Excellence Cluster, Bielefeld University,
Morgenbreede 39, 33615, Bielefeld, Germany.

Abstract

Research on situated language processing has examined how visually depicted objects or concrete action events inform the comprehension of concrete sentences. By contrast, much less is known about how *abstract* sentence comprehension interacts with non-linguistic visual information. Moreover, while non-linguistic information can rapidly inform language comprehension when it is related to sentence content through reference or lexical-semantic associations, it is unclear to which extent this is the case when the visual context is 'non-referential' (i.e., not related to the sentence through reference or lexical semantic associations). We conducted two eye-tracking reading experiments to address these two open issues. In both experiments, reading times were shorter when sentences about conceptually similar abstract ideas were preceded by objects (words-on-cards in Experiment 1 and blank playing cards in Experiment 2) that were depicted close together (vs. far apart); and when sentences about conceptually dissimilar abstract ideas were preceded by objects that were depicted far apart (vs. close together). This happened rapidly (first-pass reading times) and incrementally (as the sentence unfolded). Thus, (a) comprehension of abstract language can be modulated by non-linguistic visual information (spatial distance between depicted objects) at the sentence level, and (b) online language comprehension can be informed by visual context even in the absence of an overt referential or lexical-semantic link.

Keywords: semantic interpretation; spatial information; non-referential visual context; eye tracking.

Introduction

Studies in the 'visual world paradigm' have contributed extensively to our understanding of how non-linguistic visual information affects sentence comprehension (e.g., syntactic disambiguation: Tanenhaus et al., 1995; semantic interpretation: Sedivy et al., 1999). In 'visual world studies', listener's eye movements are tracked during comprehension of a spoken sentence that describes a given visual environment. Findings from such studies have shown that visual presentation of objects or concrete action events can facilitate incremental structural disambiguation (e.g., Tanenhaus et al., 1995; Knoeferle, Crocker, Scheepers, & Pickering, 2005); that language can rapidly guide visual attention to semantically relevant objects as evidenced by anticipatory eye-movements (e.g., Altmann & Kamide 1999; Kamide, Scheepers, & Altmann, 2003, Kamide, Altmann, & Haywood, 2003); and that distractor objects are inspected more often when they are semantically related (vs. unrelated) to a target word (e.g., Huettig & Altmann, 2005, 2011; Huettig & McQueen, 2007). Visual context not only affects spoken language comprehension rapidly, but also sentence comprehension during reading. Evidence from picture-sentence verification has revealed rapid visual context effects for concrete visual stimuli (e.g., red dots) and sentence content (e.g., *The dots are red*, see Clark & Chase, 1972; also Gough, 1965; Knoeferle, Urbach, & Kutas, 2011; Underwood, Jebbet, & Roberts, 2004).

However, most of these studies have concentrated on sentences about concrete objects and events. While evidence suggests that visual context can rapidly and incrementally inform comprehension of concrete spoken and written sentences, it is unclear to which extent non-linguistic visual context information can influence the processing of abstract language rapidly and incrementally. In examining situated language comprehension, most visual world studies have further relied on a *referential* linking hypothesis (e.g., a noun referencing an object or a verb an action). By contrast, it's unclear whether visually presented information can influence sentence comprehension when there is no overt referential or lexical-semantic link with sentence content.

Spatial Distance and Semantic Similarity

Conceptual metaphor theory proposes that abstract meaning is grounded in physical experience through metaphorical mapping (Lakoff & Johnson, 1999). Similarity, for instance, would be grounded in the physical experience of spatial distance. Recent behavioral studies have provided first evidence for a link between spatial distance and similarity. In one study, two visually presented abstract words (e.g., *loyalty* and *boredom*) were judged to be more similar when they were presented close together (vs. far apart), but more dissimilar when they were presented far apart (vs. close together, Casasanto, 2008). In another, similarity-judgment task (on whether two squares on a screen had similar colors or not) speeded decision times were shorter when similarly-colored squares were presented close to each other (vs. far apart), and when differently-colored squares were presented far apart (vs. close to each other, Boot & Pecher, 2010). These rating and response time effects support the view that there is a relationship of some sort between spatial information (the distance between two stimuli) and semantic and visual similarity.

Accounting for situated language comprehension

The nature and time course of spatial distance effects on cognitive, and in particular, language comprehension processes, however, remains unclear. In summary, we have identified several open issues in research on the interaction between non-linguistic visual information and language comprehension, which we conceptualize as two research questions: (1) Can non-linguistic information <u>rapidly</u> and <u>incrementally</u> modulate the semantic interpretation of abstract sentences? (2) Can non-linguistic visual information modulate language comprehension <u>even in the absence</u> of referential or lexical-semantic links?

Addressing these and other questions is important to advance accounts of situated language comprehension (e.g., the Coordinated Interplay Account, CIA, Knoeferle & Crocker, 2006; Knoeferle & Crocker, 2007). The CIA accommodates visual context effects during spoken language comprehension. It consists of three informationally and temporally dependent stages. A first stage accommodates the processes of incremental sentence comprehension that are the focus of traditional sentence processing accounts. A second stage describes utterance-mediated shifts in (visual) attention. 'Scene integration', finally, integrates the linguistic and scene input and informs interpretation based on visual representations. The CIA makes no assumptions regarding the modular status of either the linguistic or visual processes involved. Rather, it outlines the interaction of utterance interpretation, (visual) attention and scene information.

The CIA has been derived from eye-tracking findings on the comprehension of concrete sentences in non-linguistic visual contexts. The rapid and incremental time course with which information in visual context interacts with spoken language comprehension appears to generalize to reading when there is a referential link between visual context and sentence meaning (see Knoeferle et al., 2011 for evidence). Knowing whether rapid and incremental visual context effects are also observed when there is no referential link and when sentences are abstract, would be important for extending the account and refining its language-context linking mechanism. Based on the close time locking of visual context effects and language comprehension in the CIA, we would expect to see spatial distance effects emerge time-locked to when information about semantic similarity becomes available in the sentence. The present research addressed the two research questions (1) and (2) in two eye-tracking reading experiments. The studies examined whether, and if so with which time course, spatial distance effects on semantic similarity processing occur during comprehension of abstract sentences.

Experiment 1

In Experiment 1, participants inspected a visual context that depicted words on cards either close together or far apart (Fig. 1). Then they read a sentence that was either about similarity or dissimilarity between abstract nouns (Table 1). After reading the sentence and judging its veracity, they saw a picture and verified whether it was the same as the one that they had inspected before the reading task.

If spatial distance between the cards can modulate the interpretation of semantic similarity, we should see this reflected in reading times. To the extent that the existing findings on spatial distance effects (Boot & Pecher, 2010; Casasanto, 2008) generalize to language comprehension we should see faster reading times for similarity-conveying sentences when the preceding words-on-cards are close together (vs. far apart), and for dissimilarity-conveying sentences when the words-on-cards are far apart (vs. close together). Moreover, if effects of spatial distance are incremental, we should see them at the adjective region of the sentence (see Table 1 for examples) since this is when similarity relations are made explicit and could thus be related to spatial distance from the recent visual context. In principle, effects could appear even earlier, namely at the second noun phrase, since semantic similarity could become available as soon as the two abstract nouns are integrated. Finally, to the extent that these effects are immediate, we should observe them in first-pass reading times.

Method

Participants Thirty-two native speakers of German (mean age: 23.6; range 19-33) with normal or corrected-to-normal vision participated in the experiment for a compensation of 6 Euro. None of them had been exposed to a second language before age 6, and all gave informed consent.

Materials and Design We created 48 sentences[1], each of which had two versions. In one version the sentence expressed *similarity* between two abstract nouns, and in the other version it expressed *dissimilarity* between two abstract nouns (see Table 1 for examples). In addition, we created visual contexts using commercial graphics programs. The visual context showed two playing cards each of which presented an abstract word (see Fig. 1). Card depictions did not change between items but the words on the cards did. The words on the cards always appeared as the first and second noun phrase in the sentence. The two visual contexts and the two sentences made up an item.

A 2x2 within-subjects Latin square experimental design was implemented with two factors (spatial distance and semantic similarity), each with two levels (close vs. far, similar vs. different, respectively). Combinations of the two factors and levels resulted in four experimental conditions: cards far apart vs. close together with a similarity-conveying sentence; and cards far apart vs. close together with a dissimilarity-conveying sentence (see Table 1).

We constructed 96 filler sentences. All of them were grammatical and semantically legal German sentences. However, 72 of them described unrealistic situations (e.g.,

---

[1]One item was removed due to an error in the order of presentation of words.

'a presentation without good rhetoric should be given more often'), and the other 24 described plausible situations (e.g., 'on the tram, passengers show their ticket to the inspector'). All filler sentences were preceded by cards in different positions on the screen (e.g., in the upper left and lower right corner) and most of them were blank (N=72). Twenty-four filler sentences, however, had cards with words on them. There were in addition 14 practice trials. A list consisted of 144 trials (48 experimental and 96 fillers trials), which were all pseudo-randomized in four lists. Each list contained only one version of every item. There was at least one filler trial in between two items.
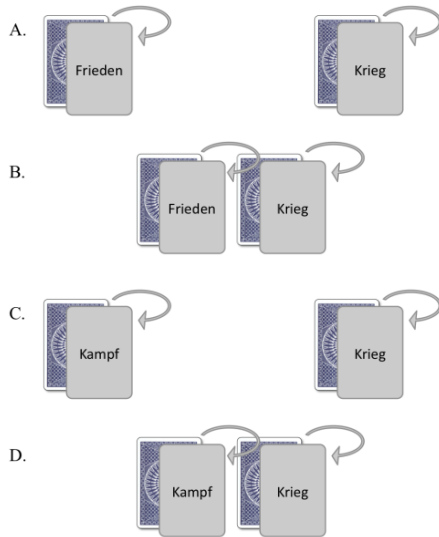

Figure 1: Example visual contexts for the sentences in Table 1 (Experiment 1). The cards moved from the center of the screen either far apart (A, C) or close to each other (B, D). After two seconds, cards turned around (as represented by the semi-circular arrow) and presented two abstract words. The words were either semantically dissimilar (e.g., *Frieden* ['Peace'] and *Krieg* ['War'], (A) and (B)) or similar (e.g., *Kampf* ['Battle'] and *Krieg* ['War'], (C) and (D)).

Procedure Upon arrival in the laboratory, participants received information about the study. After that they were calibrated using a 9-point calibration procedure. Then they completed 14 practice trials. After practice, the experiment began. Each trial was presented as a three-step task. First, participants saw a visual context for six seconds, with two playing cards in different positions. For half of the trials in the experiment (i.e., 24 filler and all experimental trials), cards turned around after two seconds, each showing a word for four seconds. For the other half of the trials, cards turned and showed a blank front for 500 ms. After inspecting the visual context participants read a sentence (see Table 1). They were instructed to try to understand the sentence and to judge its veracity (by pressing either a "yes" or a "no" button). For the ensuing picture verification, participants saw a picture of two cards, and verified (by pressing "yes"/"no" buttons) whether they were identical with the two cards they had seen before the sentence.

Table 1: Example of visual contexts, corresponding sentences, and the resulting condition.

Picture	Sentence type	Condition
Fig.1A	(1) $Frieden_{NP1}$ und_{coord} $Krieg_{NP2}$ $sind_{VP1}$ $bestimmt_{ADV}$ $verschieden_{ADJ}$,	Far-Dissimilar
Fig.1B	das $verriet_{VP2}$ der $Anthropologe_{NP3}$.	Close-Dissimilar
Fig.1C	(2) $Kampf_{NP1}$ $und_{coord.}$ $Krieg_{NP2}$ $sind_{VP1}$ $freilich_{ADV}$ $entsprechend_{ADJ}$,	Far-Similar
Fig.1D	das $verriet_{VP2}$ der $Anthropologe_{NP3}$.	Close-Similar

Translation: (1) 'Peace$_{NP1}$ and$_{coord.}$ war$_{NP2}$ are$_{VP1}$ certainly$_{ADV}$ different$_{ADJ}$, suggested$_{VP2}$ the anthropologist$_{NP3}$.' (2) 'Battle$_{NP1}$ and$_{coord.}$ war$_{NP2}$ are$_{VP1}$ surely$_{ADV}$ similar$_{ADJ}$, suggested$_{VP2}$ the anthropologist$_{NP3}$'.

Data Analysis Log-transformed reading times were analyzed using a linear mixed effect regression (LMER), including in a single step main and interaction effects of the factors. We implemented full models with random intercepts for participants and items, and fixed effect random slopes and their interactions for both random intercepts. We analyzed the second noun phrase (NP2) and the adjective region (ADJ), where we should see spatial distance effects on semantic interpretation if those occur time-locked to when semantic similarity between the first two noun phrases becomes available during reading (see Table 1). We further analyzed the VP2 and NP3 regions to see if any spatial distance effects also occur at subsequent regions of the sentence. In these regions we examined three eye-tracking reading measures; first-pass reading (the duration of all fixations from first entering an interest area and prior to moving to another interest area), regression path duration (the time from first entering a region until moving past that region to the right; unlike first-pass reading time, this measure includes reading time following regressions out of the region), and total reading times (the duration of all fixations in a given region, see, e.g., Rayner, 1998).

Results

For the ADJ region, a similarity main effect was observed across all measures; reading times were shorter for sentences expressing similarity compared to those expressing dissimilarity (all *t*-values > 2). We also observed a reliable interaction between spatial distance and semantic similarity in first-pass times for the ADJ region (*t*-value = -2.04, see Fig. 2): first-pass times were shorter when similarity-conveying sentences were preceded by cards-with-words presented close to each other (vs. far apart). By contrast, first-pass times for sentences that expressed dissimilarity were shorter when they were preceded by cards-with-words presented far apart (vs. close to each other). No interaction effects were observed in other measures. Both VP2 and NP3 regions showed reliable

interaction effect in first-pass reading times and NP3 also in total reading times. These interaction effects were similar as for the ADJ region. For the NP2 region, neither main effects nor interaction effects involving the manipulated factors were observed in any gaze measure.
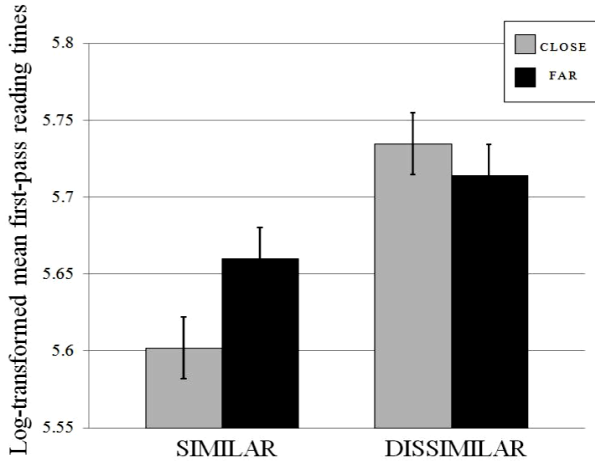

Figure 2: Log-transformed mean first-pass reading time (with error bars plotting the standard error of the mean) for the ADJ region as a function of sentence type and spatial distance between cards-with-words in Experiment 1.

Discussion

In Experiment 1, we presented participants with a visual context for which the distance between cards-with-words, was manipulated. Cards were either presented close together or far apart, and they were followed by a sentence that either expressed similarity or dissimilarity of abstract nouns. We observed spatial distance effects on reading times as a function of the semantic content of the sentence. These results suggest that non-linguistic information from the visual context (spatial distance) can modulate interpretation of abstract language (semantic similarity) during online sentence comprehension. Crucially, spatial distance effects on semantic interpretation appeared both rapidly (first-pass) and incrementally (at the ADJ region).

Results from Experiment 1, inform our first research question (whether abstract semantic interpretation can be rapidly and incrementally modulated by non-linguistic information). However, since the visual context for critical items was related to the following sentence through words on the cards, it is possible that effects of spatial distance on semantic similarity interpretation were mediated by, or even depended upon, that link. To address this concern and to answer our second research question (can non-linguistic visual information modulate language comprehension even in the absence of referential or lexical-semantic links?), Experiment 2 relies on the same design and presentation but the cards did not show any words and remained blank.

Experiment 2

Experiment 2 was identical to Experiment 1 but instead of presenting words on cards, the cards that participants saw before sentence reading were blank. Seeing the noun phrases on the cards in Experiment 1 could permit participants to integrate similarity of the nouns with spatial information even before sentence reading. If so, this could facilitate and speed up any effects of spatial distance during reading. Using blank cards in Experiment 2 permitted us to see to which extent the effects of spatial distance on sentence reading in Experiment 1 depended upon the repetition of sentential noun phrases on the cards.

Method

Participants Thirty-two further native speaker of German with normal or corrected-to-normal vision (mean age: 24.4; range 20-31) participated in the experiment for a compensation of 6 Euro. All gave informed consent.

Materials, Design, Procedure and Data Analysis The experimental design, procedure and data analysis were the same as in Experiment 1. The visual context, however, was modified. While participants in Experiment 1 saw cards-with-words for four seconds, participants in Experiment 2 saw blank cards for three seconds. In both experiments, however, visual context presentation duration was the same (six seconds). We delayed card turning by one second in Experiment 2, since participants did not have to read any words.

Results

At the NP2 region we observed a similarity main effect, such that sentences that expressed similarity had shorter first-pass times, compared to sentences that expressed dissimilarity (t-values = 2.04). Moreover, analyses of first-pass times confirmed a reliable interaction between spatial distance and semantic similarity (t-value = -2.07). Figure 3 shows the interaction pattern in Experiment 2.
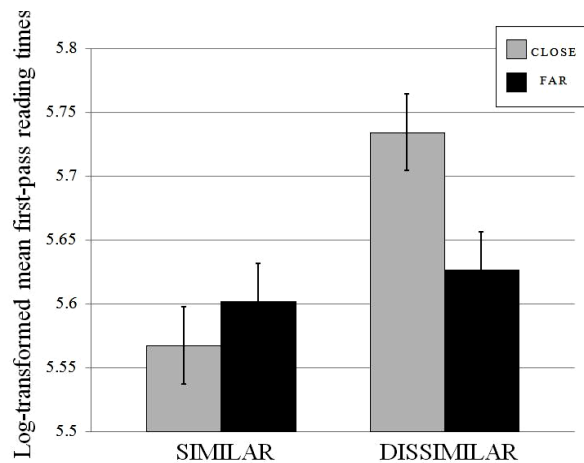

Figure 3: Log-transformed mean first-pass reading times (with error bars plotting the standard error of the mean) for the NP2 region as a function of sentence type and spatial distance between cards in Experiment 2.

First-pass times for sentences that expressed similarity were shorter when preceded by blank cards close to each

other (vs. far apart), while reading times for sentences that expressed dissimilarity were shorter when preceded by blank cards far apart (vs. close to each other). No other interaction effects were observed in other measures in this region.

Unlike in Experiment 1, a main effect of similarity appeared in regression path duration at the ADJ region, but no interaction effects were observed for that region. For the subsequent VP2 region, a marginal interaction effect emerged (t-value = -1.84), with a similar pattern to that at the NP2 region in Experiment 2 and the ADJ region in Experiment 1. Main effects of spatial distance and similarity were observed in regression path duration for the NP3 region, but no interaction effect was found.

Discussion

In Experiment 2, we manipulated the distance between objects that did not have any overt relation to the semantic meaning of the ensuing sentence. The visual context can thus be described as non-referential in its relationship to the sentence. Moreover, spatial distance between cards was irrelevant for performing the comprehension task (sentence veracity judgments). Yet, we still found rapid and incremental interaction effects between spatial distance and the interpretation of semantic similarity, as reflected in reading times. These findings provide strong support for the role of non-referential spatial information in abstract language comprehension.

General Discussion

In research on situated language processing, abstract language comprehension has received substantially less attention than concrete language. While results from visual world paradigms have shown that non-linguistic information can rapidly and incrementally modulate comprehension of sentences that relate to objects and events (through referential or associative links), it was unclear whether this effect would generalize to abstract sentences.

We assessed this open issue in two eye-tracking-reading experiments. Experiment 1 provided evidence that spatial information (distance between words depicted on cards) can rapidly and incrementally influence semantic interpretation of abstract sentences. First-pass times at the adjective and subsequent regions of sentences such as 'battle and war are surely similar, suggested the anthropologist' were shorter when a preceding display showed cards-with-words close together (vs. far apart). Since objects in the visual context were still related to the sentence in Experiment 1 (through words on cards), Experiment 2 examined whether the rapid spatial distance effects from Experiment 1 extend to a situation in which the visual context (playing cards) was entirely unrelated to the sentence. When cards remained blank rather than showing words (Experiment 2), we observed rapid and incremental effects of spatial distance on reading times as a function of the meaning of the sentences. First-pass times at the second noun phrase and at the second verb in sentences such as 'peace and war are certainly different, suggested the anthropologist' were shorter when a preceding display showed cards-with-words far apart (vs. close together).

Overall, thus, non-linguistic information can rapidly and incrementally influence semantic interpretation of abstract language. These results extend existing similarity-judgment and response-times results (Boot & Pecher 2010; Casasanto 2008), and clarify that non-linguistic information (i.e., spatial distance) can modulate language comprehension and not just similarity judgments and ratings. In this regard, our results are compatible with theories of embodied cognition that attribute an important role to perceptual information in language processing (see Lakoff & Johnson 1999). They also extend findings that suggest abstract language can be related to visual context through lexical-semantic associations. Duñabeitía et al. (2009) showed that listeners rapidly inspected objects (e.g., a nose) that were associated with abstract words (e.g., Spanish *olor*, 'smell'). What the present findings add is the insight that visual context can in turn influence abstract language comprehension, and that its effects are mediated via subtle mappings between the semantic similarity of nouns and object distance.

The present findings also have important implications for processing accounts of situated language (e.g., the CIA, Knoeferle & Crocker, 2006, 2007). Consider how concrete language is related to visual context in the CIA. The emerging interpretation guides (visual) attention to relevant information in visual context or working memory. As people process a sentential verb, they engage in a search for a matching action either in the immediate environment or in their working memory. When they have found a matching action, verb and action are co-indexed and action information can inform sentence interpretation. Comprehenders can further develop expectations about referents based on lexical-semantic associations between the verb and objects in context.

The present findings corroborate and extend the referential and associative linking mechanism of the Coordinated Interplay Account. First, the spatial distance effects occurred not anywhere during reading but they first emerged at sentence regions that contained information about semantic similarity in both experiments. This was as expected based on the CIA's close time lock between when the utterance identifies relevant visual context information, and when that context information impacts sentence interpretation. What the present results add is the insight that this closely temporally coordinated interplay extends to non-referential visual context effects (Experiment 2) and abstract language comprehension (Experiments 1 and 2). Future research will examine the subtle differences in the time course of spatial-distance effects in Experiment 1 (ADJ, NP2) compared with Experiment 2 (NP2, VP2).

The findings moreover emphasize the necessity of assuming a fine-grained linking of linguistic and visual information during language comprehension. As semantic similarity is computed during comprehension, it is reconciled with representations of spatial distance between

objects that were neither mentioned in the sentence nor relevant for the sentence judgment task. This ties in with other recent results. Kreysa and Knoeferle (2011) observed rapid visual context effects (of a speaker's gaze and head movements) on spoken language comprehension, and this despite the fact that the speaker was never explicitly referenced to and that comprehenders hardly inspected the speaker during comprehension. Clearly, explicit reference is not necessary for incremental visual context effects.

Together the present and other recent findings argue for highly active visual context effects on language comprehension and highlight the need for multiple (referential and non-referential) mechanisms in informing language comprehension through non-linguistic visual information.

Acknowledgments

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, *73*, 247-264.

Boot, I. and Pecher, D (2010). Similarity is closeness: metaphorical mapping in a conceptual task. *Quarterly Journal of Experimental Psychology, 63,* 942-954.

Casasanto, D. (2008) Similarity and proximity: When does close in space mean close in mind. *Memory and Cognition, 36,* 1047-1056.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3,* 472–517.

Duñabeitia, J. A., Avilés, A., Afonso, O., Scheepers, C., & Carreiras, M. (2009). Qualitative differences in the representation of abstract versus concrete words: evidence from the visual-world paradigm. *Cognition*, *110*, 284-292.

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of Verbal Learning & Verbal Behavior*, 4, 107–111.

Huettig, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition*, *96*, B23-32.

Huettig, F., & Altmann, G. T. M. (2011). Looking at anything that is green when hearing "frog": how object surface colour and stored object colour knowledge influence language-mediated overt attention. *Quarterly Journal of Experimental Psychology*, *64*, 122-145.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic, and shape information in language-mediated visual search. *Journal of Memory and Language, 54*, 460-482.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133-156.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*, 37-55.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: evidence from eye tracking. *Cognitive Science, 30,* 481-529.

Knoeferle, P., & Crocker, M. W. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language, 57,* 519-543.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, *95*, 95-127.

Knoeferle, P., Urbach, T. P., & Kutas, M. (2011). Comprehending how visual context influences incremental sentence processing: Insights from ERPs and picture-sentence verification. *Psychophysiology*, *48*, 495-506.

Kreysa, H., & Knoeferle, P. (2011). Effects of speaker gaze on spoken language comprehension: Task matters. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1557-62). Austin, TX: Cognitive Science Society.

Lakoff, G., & Johnson, M. (1999) *Philosophy in the flesh: The embodied mind and its challenge to western thought.* University of Chicago Press.

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

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109-147.

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.

Underwood, G., Jebbett, L., & Roberts, K. (2004). Inspecting pictures for information to verify a sentence: Eye movements in general encoding and in focused search. *The Quarterly Journal of Experimental Psychology, 56,* 165–182.