Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Is incremental semantic interpretation related to end-of-sentence verification?: Evidence from correlation analyses

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Knoeferle, Pia
Kutas, Marta
Urbach, Thomas P.

Publication Date
2009

Peer reviewed

Is incremental semantic interpretation related to end-of-sentence verification?: Evidence from correlation analyses

Pia Knoeferle (knoeferl@cit-ec.uni-bielefeld.de)
Cognitive Interaction Technology Excellence Cluster
Bielefeld University, 33615 Bielefeld, Germany

Thomas P. Urbach (turbach@cogsci.ucsd.edu)
Department of Cognitive Science
University of California San Diego La Jolla, USA

Marta Kutas (kutas@cogsci.ucsd.edu)
Department of Cognitive Science
University of California San Diego La Jolla, USA

Abstract

Prior research has typically viewed verification as a "late" process that is distinct from comprehension, occurring only after sentence comprehension is complete (e.g., Tanenhaus et al., 1976). If so, we would expect to see no clear relationship between incremental semantic interpretation processes and end-of-sentence verification times. Alternatively, these two processes may be systematically related. To examine this issue, we recorded event-related brain potentials (ERPs) as young adult participants read sentences in which the verb matched versus mismatched the action in a preceding picture. ERPs to the verb-action mismatch resembled the centro-parietal N400 typically seen in response to lexico-semantic incongruities during reading, suggesting that incremental semantic interpretation plays a key part in picture-sentence congruence processing.. Moreover, amplitudes of participants N400s to the verb correlated reliably with end-of-sentence verification latencies. Our findings contribute to the revalidation of the verification paradigm for studies of language comprehension, and provide support for the constituent-wise comparator mechanism of the Constituent Comparison Model of sentence-picture verification (Carpenter & Just, 1975).

Keywords: sentence-picture verification; depicted events; constituent comparison model

Introduction

To date, there has been relatively little research on how people reconcile what they see in a scene with a sentence that they might read (about it). The nature of such reconciliation is of interest in numerous comprehension situations such during the reading of comic books and advertisements (Carroll et al., 1992; Rayner et al., 2001), or inspection of scientific diagrams (Feeney et al., 2000). It is also relevant because linguistic utterances are often a less-than-perfect match for our current representation of the non-linguistic visual environment. To the extent that pictorial and language-derived representations are incongruous with respect to each other their reconciliation may be more or less difficult.

Questions about picture-sentence comparison processes date back to research in the 1960s (e.g., Gough, 1965). Just and Carpenter (1971), for instance, asked participants to inspect a picture of either red or black dots, followed by a related written sentence. Participants' verification latencies were found to be shorter when the color of the dots on the image (red vs. black) matched (*red*) than mismatched (*black*) the color adjective in the sentence (henceforth 'mismatch effect'). Based on these and other findings, Carpenter and Just (1975, see also Clark and Chase, 1972) developed a model of sentence-picture comparison processes (Constituent Comparison Model (CCM)). Their model accounts for response latencies in a number of sentence-picture verification studies by attributing differences to congruence (fast) versus incongruence (slow) between sentence and picture (e.g., Gough, 1965, Just & Carpenter, 1971). Findings also suggested that verification times increased linearly as a function of comparison steps, leading to the claim that verification proceeds via serial comparison of the representations of sentence and corresponding picture constituents. This also led to the claim that verification studies can provide insights into incremental sentence comprehension.

However, since reaction times in verification studies are often measured at sentence end, we cannot dismiss the possibility that they reflect verification processes that occur only after sentence comprehension. Ideally, then we need an additional continuous measure (e.g., eye tracking or event-related brain potentials, ERPs) in order to track congruence processing during real-time sentence comprehension. However, there have been relatively few sentence-picture verification studies employing either eye tracking or ERPs. One eye-tracking study reported findings that challenged the validity of the CCM, and identified additional factors (e.g., order of picture-sentence presentation) that modulated picture-sentence comparison processes (Underwood et al. 2004). In their Experiment 1, congruence (of a noun and its referent in the picture) was manipulated (match vs. mismatch); picture and caption were presented together. Analyses of both gaze data and response latencies confirmed known match-mismatch effect: Response latencies and total inspections were longer and number of fixations higher for the mismatch than for the match conditions. However, when the order of presentation (picture-first, sentence-first) was varied neither response latencies nor inspection times for the entire sentence yielded a mismatch effect. These differences were unlikely to be due to processing difficulties in resolving the mismatch, given that response accuracy was relatively high (83.6 and 79.2 percent for match and mismatch responses respectively, with no reliable difference between conditions).

Underwood et al., however, report only total sentence reading times rather than reading times for individual sentence regions or analyses of gaze measures such as first pass and regression path duration times (see Rayner, 1998 for an overview). As a consequence, such eye-tracking studies of the verification paradigm (e.g., Carroll et al., 1992; Underwood et al., 2004) have offered limited insights into congruence processing. Knoeferle and Crocker (2005) extended this work by measuring word by word as well as total sentence reading time as participants examined a pictured followed by a sentence. While total sentence reading times did not reveal mismatch effects, reading times at the word at which the mismatch became apparent (the verb and post-verbal adverb) showed clear and reliable congruence effects. Thus, with an appropriately fine-grained temporal measure, picture-sentence verification does manifest as an incremental process even during serial picture-sentence presentation. These results suggest that verification processes may occur in parallel with online language comprehension processes.

Wassenaar and Hagoort (2007) provided further evidence for this conclusion, while raising questions about the nature of the cognitive processes that contribute to verification. They compared the electrical brain activity of Brocas aphasics with that of healthy elderly adults during online thematic role assignment in a picture-sentence verification task. Participants saw a line drawing of an agent-action-patient event (e.g., man pushing woman, or a woman reading a book), and then listened to a spoken utterance in Dutch that was either a reversible active sentence (The tall man on [sic[1]] this picture pushes the young woman), a non-reversible active sentence (The young woman on [sic] this picture reads the exciting book), or a reversible passive sentence (e.g., The woman on [sic] this picture is pushed by the tall man). Healthy elderly adults exhibited a large posterior negativity (with a non-reliable late positivity) to mismatching conditions (relative to matching ones) at the verb (centro-posterior from 50-450 ms; for anterior sites from circa 50-300 ms) for semantically reversible active sentences. Irreversible active and reversible passive sentence showed an early negativity for incongruous relative to congruous trials and a subsequent (reliable) late positivity. The aphasic patients, by contrast, showed no evidence for online use of the depicted role relations at the verb. Moreover, there were no reliable differences in verification response times for either aphasic or healthy adulty, perhaps because the judgment was delayed until well after sentence completion.

These findings are consistent with other evidence for the influence of visual context information on incremental interpretation during passive comprehension (e.g., Altmann, 2004; Knoeferle et al., 2005, 2008, Knoeferle & Crocker, 2007) and act-out (e.g., Chambers et al., 2004; Sedivy et al., 1999; Tanenhaus et al., 1995) tasks. Although participants could have chosen to delay using visual context as it was often unreliable i.e., incongruous and the verification response was not required until after the end of a sentence, they did not seem to do so; rather they appeared to compare pictorial and linguistic representations incrementally. The Wassenaar and Knoeferle findings together further suggest ,at least in principle, that sentence-picture verification processes extend to serial presentation but that relevant mismatch effects may not always be apparent in response latencies and total sentence reading time data.

In sum, extant data seem to concur on incremental rather than post-comprehension picture-sentence comparison. Conclusions regarding the nature of congruence processing and its relationship to verification versus comprehension processes in these types tasks are mixed: For the reversible active and passive sentences in Wassenaar and Hagoort, ERPs at the verb had the characteristic polarity, latency, and scalp distribution of an auditory N400. For their other sentences, however, they observed an anterior negativity that was reminiscent of an N2b. A similar N2b was also observed in response to an adjective-colour mismatch (object: red square; linguistic input green square; token test), presumably reflecting the attentional detection of a mismatch rather than language processing per se (DArcy & Connolly, 1999). Vissers et al. (2008) also observed a similar anterior negativity followed by a subsequent (500-700 ms) positivity, both larger for picture-sentence mismatches (of object location) than matches.

The present study is a further opportunity - with a different type of mismatch (verb-action) - to see whether the mechanisms of picture-sentence congruence processing are a part of language comprehension (e.g., semantic interpretation) or instead invoke attentional mismatch detection. In addition, we examine the relationship between online language comprehension and end-of-sentence verification to see whether we can find evidence for or against the "verification is distinct from comprehension" position of Tanenhaus et al. (1976).

To address these issues, we recorded ERPs during word-by-word sentence reading in a picture-sentence verification task: Participants first inspected a picture and then read a sentence in rapid serial visual presentation mode (2 words/second). The sentential verb either matched the depicted action or not. If findings of Wassenaar and Hagoort extend to a verb-action mismatch, then we should get evidence for incremental processing of the verb-action mismatch during sentence reading (in the ERPs at or shortly after the verb). Such a finding would affirm the position (put forward in the Constituent Comparison Model) that picture-sentence comparison proceeds incrementally, in a constituent-wise manner.

The latency, morphology, and scalp distribution of the observed ERP effects will inform us about the nature of comprehension (and other cognitive) processes. Recall that Wassenaar and Hagoort found both comprehension-like centro-

[1]'sic' is Latin and means 'so'. It indicates that the quoted materials were reproduced verbatim from the original.

parietal N400 effects, as well as frontal negativities that resembled an N2b (linked to attentional detection of mismatches, see D'Arcy & Connolly, 1999). Accordingly, a centro-parietal N400 at the verb would provide evidence for language-based semantic interpretation of the verb-action mismatch while a relatively frontal negativity would be suggestive of the contribution of attentional mismatch detection processes and/or pictorial semantic processing (e.g., Barrett & Rugg, 1990).

In addition to delineating the time course and nature of verb-action congruence processes, we examine whether or not there is a systematic relationship between the picture-sentence congruence processing during the sentence (as reflected in average ERPs) and end-of sentence verification times. If we find such a systematic relationship (i.e., correlations), this would suggest picture sentence congruence processing during the sentence - at least for the congruence manipulation we examined - is closely related to end-of-sentence verification processes and times.

Experiment

Participants

Twenty-four students of UCSD received course credit for participation. All participants were native English speakers, right-handed, and had normal or corrected-to-normal vision.

Materials, design, and procedure We created sentences and images (using commercial graphics packages). An example image pair for one item shows a gymnast punching a journalist (Fig. 1a) and a gymnast applauding a journalist (Fig. 1b). These were paired with one of the following sentences:

(1a) The gymnast punches the journalist.
(1b) The gymnast applauds the journalist.

Figure 1: Example image

Picture-sentence pairs were pre-tested for the effectiveness of the congruence manipulation in a rating study. The resulting 160 images and sentences were used to construct 80 item sets each consisting of 2 sentences and 2 images (such as Fig. 1a and 1b, and sentences (1a) and (1b)). Together with the within-subject counterbalancing (such that each verb/action occurred once as a mismatch and once as a match; left vs. right image mirroring) the design resulted in 8 experimental lists. Each list contained an equal number of matching and mismatching experimental trials, only one occurrence of an item sentence/image, and an equal number of left-to-right and right-to-left action depictions. There were, in addition, 160 filler items in each list; half of the filler items were mismatches including full mismatch (scene and sentence were entirely unrelated), ensuring that sentence comprehension was not always contingent on the scene, noun-object reference mismatches, mismatches of the spatial layout of the scene, and mismatches of color adjectives. Moreover, the filler sentences had a variety of different syntactic structures including negation, clause-level and noun phrase coordination, as well as locally ambiguous reduced relative clause constructions.

Participants first inspected the image for a minimum of 3000ms, terminated when the participant pressed a button with their right thumb. Next, a fixation dot was presented for a random duration between 500 and 1000 ms, and followed by a sentence one word at a time, each presented for 200 ms duration with a word onset asynchrony of 500 ms. Each scene-sentence pair was followed by a pause between 500 and 1000 ms in duration. Participants indicated via a button press as quickly and accurately as possible after each sentence whether it matched (true) or did not match (false) the preceding image.

Analysis We report analyses of variance (ANOVA) on response latencies and ERPs at the verb (300-500 ms). In addition we present topographical maps of the scalp distribution of the ERP effects. Finally, we perform difference score correlations to examine the relationships between congruence effects in the N400 to the verb and at the end of the sentence in the verification times. Difference score correlations have been much discussed (Cohen & Cohen, 1983; Murray & Gonzalez, 1999). We think they are informative for our study: The difference scores provide a measure of the extent to which the processing of congruous versus incongruous trials at the verb is related to processing of these same trials immediately after sentence end. For the response latencies a positive number indicates that latencies for incongruous trials are longer than for congruous trials; a negative number means that the latencies for the incongruous trials are shorter than those for congruous trials; and zero means no difference. For ERPs, difference scores of zero also indicate no difference between incongruous and congruous ERPs; a negative number means that incongruous trials were more negative going than congruous trials (with the size of the negative number indicating the difference between congruous and incongruous trials),; and, a positive number means that the incongruous trials were more positive-going than the congruous trials.

Results

Reaction time analyses revealed reliably faster (over 150 msec) response times to the sentence-final word for congruous (1104 ms, SD = 271.72) than incongruous (1273 ms, SD = 361.07) conditions ($p < 0.01$), see Figure 2.

In addition, mismatching relative to matching trials elicited

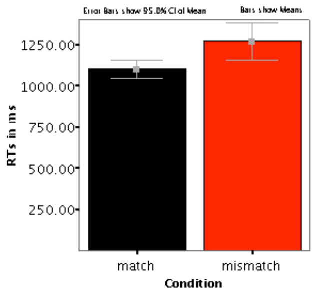

Figure 2: Response latencies in ms

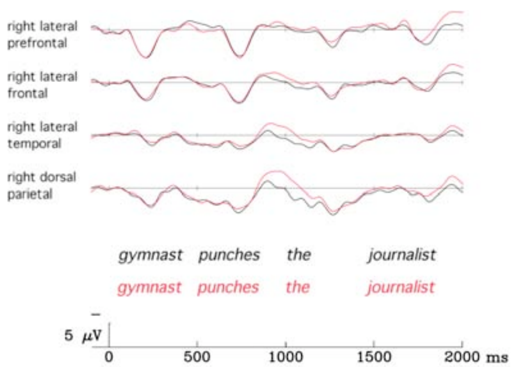

Figure 3: Verb-action mismatch N400

more negative-going event-related brain potentials (ERPs) N400s – to the verb (but not before it, $p < 0.001$). Visual inspection suggested a centro-parietal posterior maximum of the N400, as confirmed by a reliable congruence by anteriority interaction (anterior vs. posterior sites, $p < 0.05$, see Figs 3 and 4): Congruence effects were smaller over anterior (e.g., RMPf, η^2 =0.27, congruence effect n.s.) than central sites (RMCe, η^2=0.46; ps < 0.05). The overall morphology of the negativity and its amplitude distribution across the scalp resembles that of a canonical visual N400 in response to lexico-semantic incongruities (e.g., Kutas & Hillyard, 1984). An additional reliable congruence by hemisphere interaction was obtained in the 300-500 ms region of the second noun, reflecting larger effects over right than left hemispheric sites.

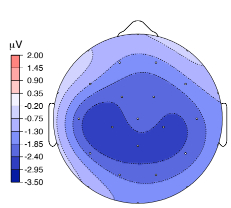

Figure 4: Scalp distribution: spline-interpolated difference scores from 300-500 ms at the verb

We found reliable correlations between each participants mismatch-match verification response time effect (mean 168.30, SE=58.59) and their N400 amplitude effect (mismatch minus match) pooled over electrodes posterior (mean N400 effect = -1.83, SD=1.83) to the vertex (r=0.53, p<0.05; r^2=0.28, see Figure 5). By contrast, this correlation was not reliable for N400s pooled over electrode sites anterior to the vertex (r=0.41 $p > 0.05$, Bonf. adjustment 0.05/2).

Inspection of Figure 5 reveals that participants with a large N400 difference at the verb (in the figure this is negative and hence small number on the x-axis) tended to have no clear congruence effect in their verification response latency (e.g., below 200 ms). As the N400 congruity effect at the verb decreased, (i.e., around zero or positive numbers on the x-axis), participants' response verification latency difference scores increased.

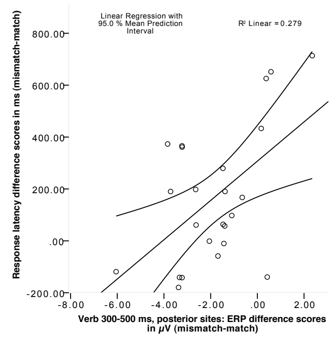

Figure 5: Scatterplots of correlations between response latency and N400 (300-500 ms) difference scores over posterior sites at the verb

General Discussion

Our results permit us to gain insight into the relationship between end-of-sentence verification (as revealed by verification response latencies) and semantic processing/comprehension (as indexed by ERPs during sentence reading). These two processes could either be unrelated, or, alternatively systematically related to one another. Our results indicate a systematic relationship between within sentence semantic processing and end-of-sentence verification times. In the following, we discuss the verification response latency findings, the ERP findings, and their relationship.

Response latency analyses showed that resolving an incongruence between a static action scene and an action verb in an immediately ensuing sentence ultimately takes more time and is presumably more difficult than when these two information sources are congruent with each other. The fact that we - unlike, for instance, Wassenaar & Hagoort, (2007), Vissers et al., (2008), and Underwood et al. (2004) - replicate the established congruence effect in response latencies (Gough, 1965; Clark & Chase, 1972; Carpenter & Just, 1975) with serial picture-sentence verification demonstrates that it was not the specific verification that led to their failures to replicate. Additional research is needed to determine why the verification response time congruence effect is sometimes present and sometimes not.

Importantly, we also find clear evidence for rapid incremental semantic interpretation (establishing reference from a verb to an action) during picture-sentence congruence processing in a serial picture-sentence verification task. Specifically, we observed larger N400s time-locked to the verb when it mismatched ('The gymnast punches') than when it matched ('The gymnast applauds') a preceding depicted action in an event scene (e.g., gymnast-applauding-journalist). If the ERP congruence effect at the verb had indexed some sort of attentional mismatch detection or pictorial processing rather than a genuine contribution of scene-based representations to sentence comprehension processes, then we likely would have obtained a frontally-distributed congruence effect similar to the N2b seen in response to a mismatch between a color adjective and the color of an object (DArcy and Connolly, 1999). This, however, was not the case. Rather, the N400 in the ERP to the verb had centro-parietal distribution, reminiscent of the N400 typically observed in response to lexico-semantic incongruities in written text (e.g., Kutas & Hillyard, 1980; Kutas & Hillyard, 1984). While the presence of an N400 effect to the verb is compatible with a lexical priming account, the subsequent ERP congruence effect (e.g., congruence by hemisphere interaction to the second noun) is not. These results are thus overall more consistent with our proposal that congruence processing - even for a lexical verb-action mismatch goes beyond lexical priming, involving more extended verification of mental representations.

Crucially, the within sentence semantic analyses effects) and post-sentence verification processes were systematically interrelated. Verb-action congruence N400 difference scores (mismatch minus match) and response latency difference scores (mismatch minus match) were reliably correlated: the larger a participant's N400 difference score, the smaller his/her verification response latency difference score. One plausible account for the correlation pattern is between-participant variation in the time course of congruence processing – participants who process the verb-action mismatch at the verb need do less verification processing later (at the end of the sentence) and hence display a smaller congruence effect in the verification response latencies). Whatever the exact account of this pattern, the correlations are clear evidence for a close relationship between end-of-sentence verification and core comprehension processes such as incremental semantic interpretation.

Findings from discourse studies using sentence-picture verification and comprehension tasks corroborate the verification-as-part-of-comprehension account. Singer (2006), for example examined whether the effect of prior discourse context on the processing of written sentences was modulated by task (answering comprehension questions vs. a combination of verification task with comprehension questions). Target sentences varied in congruence (true vs. false) with the prior discourse and negation (negated vs not negated). Singer's reading time data replicated a key finding in the sentence-picture verification literature: true negatives were harder than false negatives, regardless of task. Ferretti, Singer, and Patterson (in press) extended these findings using ERPs with these materials in a reading comprehension task (answering yes / no comprehension questions). They found congruence effects in both early (P2b) and later (late phase of the centro-parietal N400) ERP components. Taken together with our findings, it seems that the verification paradigm can play an important role in a broad range of studies on language comprehension with strictly language or visual scene contexts.

In this respect, it will be interesting to see to what extent our results and interpretations - that verification is part of situated comprehension, and that people continually verify linguistic and pictorial representations - generalize to other paradigms (e.g., 'visual worlds') and situated spoken comprehension. Moreover, it is important to discover how operations such as 'verification' relate to processes of establishing reference from a word to an object, processes of visually anticipating objects, and visual context influences on language comprehension that have been observed in the visual world paradigm. Knoeferle and Crocker (2007) propose that jointly with referential and anticipatory search of utterance-relevant objects, the comprehensions system reconciles the current and expected linguistic interpretation with scene-based mental representations by indexing nouns/verbs with objects/actions (or representations thereof), and by revising - when a mismatch is detected - the interpretation based on the scene representations. If our findings generalize, we should find evidence for a congruence effect in gaze data, at the moment when a mismatch is detected. We should also observe between-participant variation in the time course of congruence processing. Such data would further corroborate our conclusion that verification is part of online language comprehension.

One argument against a close relationship between verification and comprehension is to say that verification processes are not part of "normal" comprehension. Under this account verification is a"special" and rare case, and clearly distinct from routine comprehension processes. However, this simply is not the case. We often utter statements that verify facts in everyday life: Positive verification, for example, may be inferred from expressions of agreement "So I heard", "No doubt") while failures to verify may be inferred from corrections and requests for clarification and the like (e.g., "Well no, actually what happened was ...", "Are you sure?"). Thus, while a button press indicating 'True or 'False in a verification task may be a somewhat unnatural laboratory proxy, the verification processes and the generation of an overt response are clearly a part of routine language communication.

To be clear however, we do not maintain that our findings show that verification response times are solely a function of comprehension difficulty at the verb or that they reflect all aspects of comprehension. Sentence comprehension involves complex inferential processes that may or may not be reflected in verification latencies and/or tasks. Further-more, other decision- and response-related processes may be involved and these presumably may contribute to verification response times differentially in matching and mismatching conditions. However, if verification times reflected only those processes that are downstream and distinct from comprehension processes, we would not have expected to see a systematic relation as we did - between verification times and verb-action congruence effects. In short, our findings are consistent with a constituent-wise comparator mechanism that supports incremental picture-sentence comprehension.

Acknowledgments

This research was conducted at UC San Diego, funded by a postdoctoral fellowship to PK (DFG) and by NIH grants HD-22614 and AG-08313 to MK. We thank three reviewers for their helpful comments.

References

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

Barrett, S. & Rugg, M. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition 14*, 201-212.

Carpenter, P. A. & Just, M. A. (1975). Sentence comprehension: a psycholinguistic processing model of verification. *Psychological Review*, *82*, 45-73.

Chambers, C.G., Tanenhaus, M., & Magnuson, J. (2004). Actions and affordances in syntactic ambiguity resolution. *JEP:LMC, 30*, 687-696.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cog Psy, 3*, 472-517.

Carrol, P. J., Young, J. R., & Guertin, M.S. (1992). Visual analysis of cartoons: A view from the far side. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 444-461). NewYork: Springer-Verlag.

Feeney, A., Hola, A. K. W., Liversedge, S. P., Findlay, J. M., & Metcalf, R. (2000). How people extract information from graphs: Evidence from a sentence-graph verification paradigm. In M. Anderson, P. Cheng, & V. Haarslev (Eds.), *Theory and application of diagrams: Diagrams 2000* (pp.149-161). Berlin: Springer-Verlag.

Ferretti, T.R., Singer, W., & Patterson, C. (2008). Electrophysiological evidence for the time course of verifying text ideas. *Cognition, 108*, 881-888.

Goolkasian, P. (1996). Picture-word differences in a sentence verification task. *Memory & Cognition*, *24*, 584-594.

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *J of Verbal Learning and Verbal Behavior*, *5*, 107-111.

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with qualification. *Journal of Verbal Learning and Verbal Behavior*, *10*, 244-253.

Knoeferle, P. and Crocker, M. (2005). Incremental Effects of Mismatch during Picture-Sentence Integration. In: *Proceedings of the 26th Cogsci Conference*, Stresa, Italy.

Knoeferle, Matthew Crocker, Martin Pickering, & Christoph Scheepers (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. & Crocker, M.W. (2007). The influence of recent scene events on spoken comprehension: evidence from eye-movements. *JML (Special issue: Language-Vision Interaction), 57*, 519-543.

Knoeferle, P., Habets, B., Crocker, M.W., & Muente, T. (2008). Visual scenes trigger immediate syntactic reanalysis: evidence from ERPs during situated spoken comprehension. *Cerebral Cortex, 18*, 789-795.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

Kutas, M. & Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature 307*, 161163.

Luedtke, J., Friedrich, C.K., de Fillipis, M., & Kaup, B. (2008). Event-related potential correlates in a sentence-picture verification paradigm. *J Cog Neurosci, 20*, 1355-1370.

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

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *JEP Applied*, *7*, 219-226.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109-148.

Singer, M. (2006). Verification of text ideas during reading. *JML, 54*, 574-591.

Tanenhaus, M. K., Carroll, J. M., & Bever, T. G. (1976). Sentence-picture verification models as theories of sentence comprehension: A critique of Carpenter and Just. *Psychological Review, 83*, 310-317.

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. *QJEP*, *56*, 165-182.

Vissers, C., Kolk, H., van Meerendonk, N., & Chwilla, D. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, 46, 967-982.