

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Wiggle, Wiggle, Wiggle: How Visual Cues Influence Thematic Role Assignment in Children and Adults

Permalink

<https://escholarship.org/uc/item/32f000fb>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

Authors

Kroger, Julia Marina

Munster, Katja

Burigo, Michele

et al.

Publication Date

2018

Wiggle, Wiggle, Wiggle: How Visual Cues Influence Thematic Role Assignment in Children and Adults

Julia Marina Kröger (jkroeger@cit-ec.uni-bielefeld.de)¹

Katja Münster (muenstek@hu-berlin.de)²

Michele Burigo (mburigo@cit-ec.uni-bielefeld.de)¹

Pia Knoeferle (pia.knoeferle@hu-berlin.de)²

¹ Center of Excellence Cognitive Interaction Technology (CITEC), Bielefeld, Germany

² Department of German Studies and Linguistics, Berlin, Germany

Abstract

German 5-year-olds are able to rapidly recruit depicted actions to assign thematic roles in unambiguous sentences when these actions can be inspected throughout sentence presentation (Münster, 2016; Zhang & Knoeferle, 2012). In two visual-world eye tracking studies, we investigated whether these findings extend to locally structurally ambiguous utterances and to short-lived action presentation. In addition, we compared the action depiction to a character's wiggling motion. The action and the wiggle served as cues to the agent (subject) in difficult-to-understand OVS sentences. Participants listened to structurally ambiguous object-verb-subject (OVS) sentences about, for instance, a bug being pushed by a bull while inspecting a bull, a bug, and a worm. We manipulated the scene at verb-onset such that either a) no action no wiggle, b) no action one wiggle, c) one action no wiggle, or d) one action one wiggle appeared. Both of these animations caused the adults and the children to visually anticipate the agent role filler (corresponding to the subject in the OVS sentence) before its mention. However, in answering post-trial who-does-what-to-whom comprehension questions, the children did not (unlike suggested by previous findings) benefit from the action depictions. Together the eye-gaze and post-trial comprehension results suggest that the nature of cue presentation (e.g., the abrupt onset of an action or a wiggle and limitations on cue presence) plays an important role in both the immediate visual attention and somewhat later interpretation effects of such visual cues during children's language comprehension.

Keywords: Visual-world paradigm, eye movements, child language comprehension, thematic role assignment, depicted actions, wiggle, non-linguistic visual cues

Introduction

Adult comprehenders can exploit a variety of non-linguistic cues in their incremental interpretation of spoken utterances. Visual referential context, contrast between objects, object affordances, depicted actions, or events can each rapidly influence spoken language comprehension (e.g., Chambers, Tanenhaus, & Magnuson, 2004; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). The utterance, in turn, can guide visual attention to objects in the scene, and that attention can be 'anticipatory' such that adult listeners visually anticipated a referent even before it was mentioned (Altmann & Kamide, 1999).

For children, however, some but not all of these aspects of situated language (processing) are present from day one. First evidence for the close temporal coordination of visual attention and auditory input comes from a study on 6-month-olds

(Richardson & Kirkham, 2004). These experienced an object (e.g., a toy) jointly with a sound (e.g., *boing boing*) and immediately after listened to a spatially non-informative *boing* sound without the toy present. The infants inspected the side of the screen that the associated toy had just occupied more than the other side of the screen. Similarities in children's and adults' visual attention behaviour also emerged during utterance comprehension. Much like adults, 10-11-year-olds (Nation, Marshall, & Altmann, 2003) and 2-year-olds (Mani & Huettig, 2012) rapidly anticipated an upcoming target (a cake vs. a bird as a distractor object) during *eats the* when they listened to *The boy eats the big cake* (vs. *The boy sees the big cake*); but at the age of two this was only the case for skilled (vs. unskilled) language producers (Mani & Huettig, 2012). Eye-movements in this paradigm also revealed differences between child and adult language processing. For instance, while 36-month-olds much like adults rapidly gazed at a blue car when responding to questions such as '*Can you find the blue car?*' in the presence of a red and a blue car, 30-month-olds shifted their gaze more often towards an incorrect referent (Fernald, Thorpe, & Marchman, 2010).

Further differences emerged in referential context effects during structural disambiguation (Trueswell, Sekerina, Hill, & Logrip, 1999). Five-year-olds listened to ambiguous (e.g., *Put the frog on the napkin in the box*) and unambiguous (e.g., *Put the frog that's on the napkin in the box*) instructions and inspected contexts which contained either one or two possible referents for *the frog* (one-referent context: a frog, a horse, an empty napkin, a box; two-referent context: a frog on a napkin, another frog, an empty napkin, and a box). In the ambiguous instruction, *on the napkin* could either be interpreted as the location or as the destination of *the frog*. The 5-year-olds interpreted the *napkin* as the *frog's* destination even when the two-referent context biased towards a location interpretation. Unlike adults, the five-year-olds were unable to infer from the referential context (two frogs) that *on the napkin* modified *the frog*. These and related findings support the idea that child language processing differs from adult language processing.

However, children's thematic role assignment in structurally unambiguous German sentences was influenced by depicted action events (Münster, 2016; Zhang & Knoeferle,

2012). While listening to unambiguous subject-verb-object (SVO: e.g., *Der Bär schubst sogleich den Stier.* - ‘The bear (subj) pushes immediately the bull (obj).’) and object-verb-subject (OVS: e.g., *Den Bär malt sogleich der Wurm.* - ‘The bear (obj) paints immediately the worm (subj).’) sentences, 4-5-year-olds inspected scenes containing a bear, a worm, and a bull. Actions were (vs. weren’t) depicted (e.g., for the verb *push* the bear either held its hands towards the bull (vs. close to its own body). When actions were depicted, children visually anticipated the patient (i.e., the worm vs. agent: the bull) in SVO sentences before its mention (during *sogleich*, ‘soon’) and the agent (vs. patient) in OVS sentences (compared with when the actions were not depicted). The results for the adult control group were similar but adults started to inspect the target agent one word earlier (during the verb). Depicted actions further improved children’s (but not adults’) responses to post-sentence comprehension ‘who does what to whom’ questions for the OVS sentences. It seems children can rapidly recruit depicted action events at least for interpreting structurally unambiguous utterances. Adults can even recruit them for incremental thematic role assignment in locally structurally *ambiguous* SVO and OVS sentences (Knoeferle et al., 2005).

Prior research has further investigated the effects of different types of information in scenes on language comprehension: speaker gaze vs. a dot cursor (Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008), speaker gaze vs. an arrow (Staudte, Crocker, Heloir, & Kipp, 2014), actor gaze vs. recent action events (Abashidze, Knoeferle, & Carminati, 2015), depicted actions vs. a recent emotional speaker face (Münster, Carminati, & Knoeferle, 2015), and speaker gaze shifts vs. depicted actions (Kreysa, Knoeferle, & Nunneman, 2014). Although all of these visual cues rapidly informed language comprehension in adults, some had similar whereas others had distinct effects on language comprehension (e.g., depicted actions and speaker gaze shifts elicited looks to a target character with the same time course when they were both employed as deictic cues; the time courses differed when the verb had to be processed more in-depth semantically; an emotional prime face only seemed to affect the listeners’ visual attention when an action was also depicted). The extent to which distinct visual cues can guide comprehenders’ visual attention and influence their language comprehension is largely an open question, as is the extent to which they are processed in a similar fashion by children and by adults for disambiguating non-canonical sentences. The latter are difficult to process due to an SVO over OVS word order bias in German. It is further unclear to which extent these cues influence thematic role assignment when they (unlike in prior research) appear for only a short time during comprehension, similar to some linguistic cues.

The Present Research

The present research, thus, assessed whether depicted action events (that are mediated by the verb) rapidly influence children’s thematic role assignment in locally structurally am-

biguous OVS sentences when the action is only co-present for a short period of time. We explored the influence of further distinct cues (a wiggling motion of a character). A wiggling motion presents an interesting comparison since it is not mediated by the verb but co-located with the agent (and might thus attract the listeners’ attention to the agent). The wiggle could function as a pragmatic/focusing cue to the extent that comprehenders infer that the wiggling (vs. not wiggling) character is the agent, disambiguating sentence structure. Does a wiggling target character help children to correctly assign thematic roles (eliciting agent inspection) or do they fail to draw pragmatic inferences as suggested by Trueswell et al. (1999) where children had to infer that the napkin modified one of the two frogs? If not per pragmatic inferences, a wiggling character might first attract the comprehenders’ attention through the abrupt motion, resembling an action, implicating thematic role assignment processes.

The participants inspected a scene (Fig. 1) and listened to OVS sentences (Fig. 2). Figure 2 depicts cue presentation. If children and adults rely on distinct short-lived visual cues for correct thematic role assignment, they should look more to the target character (the agent: the bull) during or shortly after they have heard the verb when a visual cue had been present (vs. absent) during the verb. If one cue is stronger than the other, we expect a difference in looks to the target character for the depicted action and the wiggle condition. If the addition of visual cues (depicted action plus wiggle) has beneficial effects, we expect a preference to inspect the agent in the depicted action plus wiggle condition compared to the depicted action or wiggle (one-cue) condition.

In addition to eye gaze we measured comprehension via post-sentence comprehension questions. We expected to find a difference between children and adults. In ambiguously case marked OVS sentences, case marking on the second noun phrase disambiguates the sentences. Prior research has suggested that adults can use case marking for correct thematic role assignment (Kamide, Scheepers, & Altmann, 2003; Matzke, Mai, Nager, Rüsseler, & Münte, 2002; Zhang & Knoeferle, 2012). If so, adults should make correct responses independent of cue presence. Children’s ability to use case marking (in the absence of helpful visual context), however, seems to be limited to visual contexts that provide further information such as world knowledge (Özge, Münster, Knoeferle, Küntay, & Snedeker, 2016 but Kröger, Münster, & Knoeferle, 2017; Dittmar, Abbot-Smith, Lieven, & Tomasello, 2008; Schipke, Friederici, & Oberecker, 2011). Our visual contexts did not provide any such further information, and children should be at chance in responding in the no-cue condition. If children are, however, able to use visual cues for the correct interpretation of ambiguous OVS sentences, we should replicate the findings from Zhang and Knoeferle (2012): Improved accuracy when actions are depicted (vs. not depicted), perhaps even more in passive voice comprehension questions (Münster, 2016). If the wiggle effects resemble action effects, we expect more correct responses when

the wiggle is present (vs. absent). If one cue is stronger than the other, we should observe a difference in the amount of correct responses between the action and the wiggle cue. If two visual cues are better than one, we expect more correct responses in condition d) one action one wiggle compared to the single-cue conditions b) and c)).

Experiment

Participants

24 young monolingual German adults (mean age = 27.8) and 24 monolingual five-year old German kindergarten children participated in the experiment. All of them had normal or corrected vision and hearing. Adults received 5 Euros and children received a toy and a certificate. Adults and the children’s parents gave written informed consent and the children gave oral informed consent. The ethics committee of Bielefeld University approved the experiment.

Materials

A linguistically trained female native German speaker recorded 24 ambiguously case marked transitive German object-verb-subject sentences. All sentences were ambiguously case marked on the first noun phrase (i.e., either feminine or neuter case - identical in nominative and accusative) and were assigned an OVS biasing prosodic contour (L*+H accent on the first noun phrase). Prosody was not part of the design, and we did not expect to find effects of prosody (see Kröger et al., 2017). Case marking on the second noun phrase was unambiguous and thus disambiguated who does what to whom. For each sentence, we created four visual scenes containing three clipart animal characters. In condition a) no action no wiggle, the scene contained the role fillers only (Fig. 1). In condition b) no action one wiggle, the agent wiggles up and down a fixed number of pixels. Condition c) one action no wiggle depicts the agent performing an action (e.g., for the verb *schubsen* ‘push’ the character was holding his hands out). In condition d) one action one wiggle the agent performs an action and wiggles at the same time (Table 1). All cues were only present during the verb.

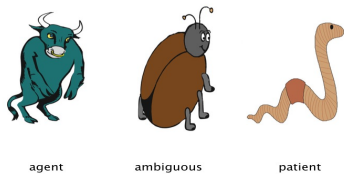


Figure 1: Example image: condition a) no action no wiggle.

The middle character was always mentioned at the beginning of the sentence and was thus role ambiguous: it could act upon or being acted upon by one of the adjacent characters (the NP1 was ambiguously case marked and did thus not disambiguate who does what to whom). In Figure 1 the bull is the agent of the scene and the worm the patient. We counterbalanced the role fillers such that in another scene the worm

is the agent and the bull the patient. We avoided stereotypicality (a bull is not a more stereotypical pusher of a bug than a bug in relation to a worm). The scenes were counterbalanced for left and right direction. In an additional 72 filler items we varied the number of depicted role fillers such that either one, two, or three role fillers were depicted and we used different sentence structures (SVOO, directOVSO, indirectOVSO) and different adverbs (manner, frequency, place). To avoid a clear OVS-biasing prosodic structure, half of the filler items were assigned an OVS-biasing prosody (L*+H on NP1) and half an SVO-biasing prosody (L*+H on NP1, H* on verb). One half of the filler items occurred in one of the experimental conditions and in the other half actions unrelated to the verb were depicted and characters other than the agent wiggled. For children, filler items were reduced to 8.

Table 1: Conditions

Condition	Sentence Structure	Visual Cue
a	ambOVS	no action no wiggle
b	ambOVS	no action one wiggle
c	ambOVS	one action no wiggle
d	ambOVS	one action one wiggle

Procedure

During the experiment, we monitored participant’s eye movements with an Eyelink 1000 eye tracker. In the remote setup with a 16mm lens, the sampling rate was 500 Hz monocular and the average accuracy 0.5°. Scenes were presented on a DELL laptop (screen resolution 1920x1080 pixels). Before the experiment, the experimenter conducted a manual calibration using a five-dot scheme. Participants saw the scene 2000 ms before sentence onset. The comprehension question, either in active or in passive voice (active: *Wer schubst hier?* ‘Who pushes here?’; passive: *Wer wird hier geschubst* ‘Who is being pushed here?’), followed 1500 ms after sentence offset. All visual cues were time-locked to verb onset (Figure 2) because at this point in time the verb referenced the depicted action. The wiggle also occurred at verb onset so that we could perform a direct comparison between the different cues. The trials were separated by a drift correct point to verify calibration. Before the experimental items and the final calibration, practice items (N=4) introduced the experiment. The duration of each testing session was approximately 40 minutes for adults and 20 minutes for children.

Analysis

For the analysis, we predefined two word regions of interest: verb (verb onset to adverb onset) and adverb region (adverb onset to NP2 onset). We were interested in the verb region to observe the effects of the visual cues. We chose the adverb region for post-verbal effects. Within each scene, we predefined two areas of interest: the left and right role filler (e.g., the bull - the agent and the worm - the patient). We did not include the middle role filler in the analyses because our dependent measure was the anticipation of the NP2 role fillers (see

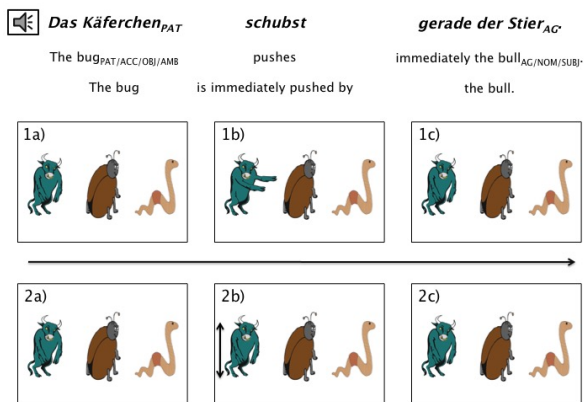


Figure 2: Visual cue presentation: Depicted actions: 1a) no cue 1b) cue depiction 1c) no cue. Wiggle: 2a) no cue, 2b) cue depiction (target character wiggles up and down 10 pixels - the arrow is for visualisation only), 1c) no cue.

also Knoeferle et al., 2005). For the two time windows, we computed mean-log ratios of looks (see Arai, Van Gompel, & Scheepers, 2007; Carminati & Knoeferle, 2013). These log-ratios represented the preference of looks towards the agent over the patient ($\ln(\text{agent})/\ln(\text{patient})$). Negative numbers indicate more looks towards the patient (vs. agent) and positive numbers more looks towards the agent (vs. patient). The mean-log ratios of looks were subjected to an analysis of variance (ANOVA) by subjects and by items with a 2 (action) x 2 (wiggle) design. Post-sentence comprehension questions include correct (1) and false/no (0) responses. Based on the number of possible responses, we calculated the percentages of corrects responses by condition. We analysed the accuracy data using Generalised-Mixed-Effects Models (Bates, Mächler, Bolker, & Walker, 2014).

Results

Eye Movements. The analyses revealed main effects of action and wiggle and an interaction of action and wiggle during the verb and adverb region (all $ps < .01$) in both age groups (Figures 3, 4). In both word regions and age groups, paired-sample t-tests after Bonferroni (.05/6) revealed significant differences for the no-cue condition versus each of the visual cue conditions (verb, adverb: $ps < .01$). Adults and children preferred to look at the agent more when one or two of the cues were present compared to when no cue was. Descriptively, the adults' mean-log ratios were higher for the verb than adverb region, indicating a more pronounced preference to look at the agent (vs. patient) during the verb than adverb. For children, paired-sample t-tests after Bonferroni (.05/6) showed a significant difference between condition b) no action one wiggle and condition d) one action one wiggle ($ps < .01$) during the adverb. Children's preference to look at the agent (vs. patient) was boosted by the wiggle in the no-action but not the action conditions. In children, the mean-log ratios were higher during the adverb than verb region.

Accuracy. For adults the overall accuracy was 68.9%. The analyses of their scores revealed a marginal interaction ($p = .07$; the wiggle modulated accuracy in the action present but not the action-absent conditions; one action one wiggle 75% vs. the other conditions (a) 66.7%, b) 66%, c) 68.1%). Children's accuracy was 41.7%. An effect of voice reflected that children responded more accurately to active than passive voice questions ($p < .001$). The active voice data revealed a marginal effect of wiggle ($p = .09$). The effect of action was neither reliable overall nor in the active-question data. Children responded more accurately to active-voice questions when the wiggle was present (vs. absent).

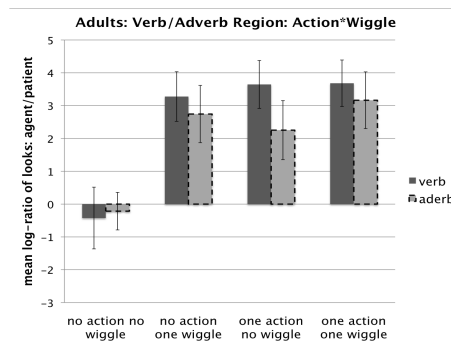


Figure 3: Adults: Mean log-ratios of looks towards the agent over the patient during the verb and adverb region in all four conditions. Error bars: 95% confidence intervals.

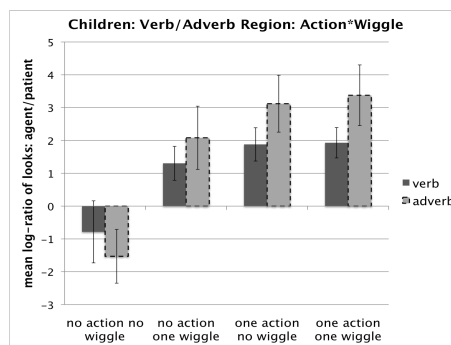


Figure 4: Children: Mean log-ratios of looks towards the agent over the patient during the verb and adverb region in all four conditions. Error bars: 95% confidence intervals.

Discussion

We investigated the influence of distinct short-lived visual cues on thematic role assignment in ambiguously case marked German OVS sentences in children and adults.

Eye-movements. The results corroborate the rapid depicted action effects reported by Zhang and Knoeferle (2012) and Münster (2016). But unlike in Zhang and Knoeferle (2012) and Münster (2016), looks to the target character in children were not delayed by one word region in our experiment. We did observe some delay, however: Descriptively, for adults the presence of visual cues resulted in a higher preference to look at the agent during the verb (vs. adverb)

whereas children's preference was higher in the adverb than verb region. The difference in the time course of gaze patterns likely resulted from presentation differences. In our study, the action and wiggle cues were limited to the verb, perhaps eliciting children's more immediate attentional response, while they had been present throughout the sentence in prior research. In addition to the action effects, a main effect of wiggle emerged, as well as an interaction of action and wiggle, all during the verb. While the action-based anticipation of the target agent was not modulated by wiggle presence in the children, when no action was present, children inspected the target agent more when the wiggle was present vs. absent.

Accuracy. For adults, a marginal interaction emerged: When the action was present, an added wiggle (vs. no wiggle) elicited more correct responses; by contrast, wiggle presence did not influence accuracy when no action was present. Surprisingly, adult's overall accuracy was very low (68.9%) although case marking on the second noun phrase disambiguated the sentences. The low accuracy on critical items may have resulted from language-scene mismatches in the fillers. These may have led the adults to interpret the case marking on the second noun phrase of OVS sentences as a 'mismatch', thinking it must be SVO, resulting in incorrect role assignment and responses to the questions.

For children, we did not replicate the improved accuracy (with vs. without action depiction) reported previously (Zhang & Knoeferle, 2012). Unlike in the adult data, we found an effect of voice in the child accuracy data such that children responded significantly more often correctly to questions in active than passive voice. Children might have used the relationship between the visual cue and the agent: The correct response to active voice questions was the agent. Since the visual cue was either the agent performing an action or the agent wiggling, children may have used this relationship in responding to active voice questions. Follow-up analyses on the active-question data revealed a marginal effect of wiggle such that the wiggle had a positive effect on accuracies for active voice questions. Perhaps the children used the wiggle to keep the agent representation in working memory, facilitating access in response to the questions.

Our eye-gaze results, however, do not corroborate the idea that the wiggle in particular boosts attention. In the adverb region, children looked significantly more often to the agent (vs. patient) when one action (vs. one wiggle) was depicted. Perhaps the wiggle increased the salience and / or focus of the agent representation in children with some delay only, eliciting increased accuracy when the agent was the question target but not eliciting more inspection in real time. Future experiments could assess this interpretation.

Why then did the depicted action not boost children's responses to comprehension questions as was the case in Münster (2016); Zhang and Knoeferle (2012)? One of the reasons why we failed to observe beneficial effects of depicted actions might be that our actions were depicted only for a short period of time. A wiggling target character could, however, also in-

terfere with thematic role assignment. Children acquire verb-argument structure and associated abstract knowledge from an early age (e.g., Bencini & Valian, 2008; Messenger, Branigan, McLean, & Sorace, 2012; Peter, Chang, Pine, Blything, & Rowland, 2015). Upon hearing a known transitive verb, children may know that the verb requires two arguments (e.g., agent and patient). The depicted action does indeed represent a two-argument event (e.g., the bug pushes the bull) and is thus compatible with the argument structure of the verb. The wiggle could, however, introduce a one-argument event (e.g., the bull wiggles) and thus be incompatible with the argument structure of the verb, perhaps reducing its immediate effects on eye gaze compared with the action depiction. But the wiggle effects on question accuracy suggests that participants at least by sentence-end had integrated the wiggle into the argument structure of the verb, perhaps because it had been - in its presentation - time-locked to the verb.

In summary, our results support the idea that children's and adults' visual attention is immediately (at the verb) guided by the visual cues. But children's post-trial response accuracy differed from previous findings and differed from that of the adults. Children - unlike in prior research seemed to be affected more by the the wiggle in their responses while adults' accuracy was highest for the combination of the two visual cues. Perhaps the temporal limitations in the presentation of the cues resulted in an immediate boost but with the negative effect that the children processed the actions less in-depth, eliminating the previously-observed benefit on their accuracy in responding to who-does-what-to-whom questions. Since children around the age of five still seem to struggle in interpreting ambiguous and non-canonical sentence structures (SVO preference), the use of visual support (even if short-lived) may be helpful to encourage language development.

Acknowledgments

This research was funded by the Project "FoTeRo" in the Focus center XPrag (DFG) and by the Cognitive Interaction Technology Excellence Cluster (277, DFG). We thank Franziska and Doris Müller for their help in recruiting child participants and finding testing locations at kindergartens.

References

- Abashidze, D., Knoeferle, P., & Carminati, M. N. (2015). Eye-tracking situated language comprehension: Immediate actor gaze versus recent action events. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 31–36).
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.
- Arai, M., Van Gompel, R. P., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology*, 54(3), 218–250.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

- Bencini, G. M., & Valian, V. V. (2008). Abstract sentence representations in 3-year-olds: Evidence from language production and comprehension. *Journal of Memory and Language*, 59(1), 97–113.
- Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, 106(3), 1465–1477.
- Carminati, M. N., & Knoeferle, P. (2013). Effects of speaker emotional facial expression and listener age on incremental sentence processing. *PLoS ONE*, 8(9), e72559.
- Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(3), 687–696.
- Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development*, 79(4), 1152–1167.
- Fernald, A., Thorpe, K., & Marchman, V. A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective–noun phrases. *Cognitive Psychology*, 60(3), 190–217.
- Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32(1), 37–55.
- 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(1), 95–127.
- Kreysa, H., Knoeferle, P., & Nunneman, E. M. (2014). Effects of speaker gaze versus depicted actions on visual attention during sentence comprehension. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (pp. 2513–2518).
- Kröger, J. M., Münster, K., & Knoeferle, P. (2017). The influence of prosody and case marking on thematic role assignment in ambiguous action scenes: Adults versus children. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 2463–2468).
- Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, 38(4), 843–847.
- Matzke, M., Mai, H., Nager, W., Rüsseler, J., & Münte, T. (2002). The costs of freedom: an ERP-study of non-canonical sentences. *Clinical Neurophysiology*, 113(6), 844–852.
- Messenger, K., Branigan, H. P., McLean, J. F., & Sorace, A. (2012). Is young children's passive syntax semantically constrained? Evidence from syntactic priming. *Journal of Memory and Language*, 66(4), 568–587.
- Münster, K. (2016). Effects of emotional facial expressions and depicted actions on situated language processing across the lifespan. *Doctoral Thesis, Bielefeld University*.
- Münster, K., Carminati, M. N., & Knoeferle, P. (2015). The effect of facial emotion and action depiction on situated language processing. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (pp. 1673–1678).
- Nation, K., Marshall, C. M., & Altmann, G. T. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, 86(4), 314–329.
- Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, 81, 1–15.
- Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General*, 133(1), 46.
- Schipke, C. S., Friederici, A. D., & Oberecker, R. (2011). Brain responses to case-marking violations in German preschool children. *Neuroreport*, 22(16), 850–854.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–147.
- Staudte, M., Crocker, M. W., Heloir, A., & Kipp, M. (2014). The influence of speaker gaze on listener comprehension: Contrasting visual versus intentional accounts. *Cognition*, 133(1), 317–328.
- 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*, 1632–1634.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying online sentence processing in young children. *Cognition*, 73(2), 89–134.
- Zhang, L., & Knoeferle, P. (2012). Visual context effects on thematic role assignment in children versus adults: Evidence from eye tracking in German. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 2593–2598).
- Özge, D., Münster, K., Knoeferle, P., Küntay, A., & Snedeker, J. (2016). Predictive use of German case markers in German children. In *Proceedings of the 40th Annual Boston University Conference on Language Development* (pp. 291–303).