Visual Context Effects on Thematic Role Assignment in Children versus Adults: Evidence from Eye Tracking in German

Lu Zhang (lzhang@cit-ec.uni-bielefeld.de)
Pia Knoeferle (knoeferl@cit-ec.uni-bielefeld.de)
Cognitive Interaction Technology Excellence Center, University of Bielefeld,
Bielefeld, Germany

Abstract

Prior research has shown that adults can make rapid use of visual context information (e.g., visual referential contrast and depicted agent-action-patient events) for syntactic structuring and disambiguation. By contrast, little is known about how visual context influences children's language comprehension, and some results even suggest children cannot use visual referential context for syntactic structuring (e.g., Trueswell et al., 1999). We examined whether children (unlike adults) also struggle to use other kinds of information in visual context (e.g., depicted events) for real-time language comprehension. In two eye-tracking studies we directly compared real-time effects of depicted events on children's (Exp1) vs. adults' (Exp2) processing of spoken German subject-verb-object (SVO) and object-verb-subject (OVS) sentences. Both of these word orders are grammatical, but OVS is a non-canonical structure. Five-year olds are at chance in understanding even unambiguous OVS sentences in the absence of visual context (Dittmar et al., 2008). If children can use depicted events rapidly for syntactic structuring, we should find similar visual context effects for them as have been reported for adults (Knoeferle et al., 2005), and similar gaze pattern as for the adults in the present studies. Gaze pattern in the present studies suggested that events depicting who-does-what-to-whom incrementally influenced both adults' and 5-year-olds' visual attention and thematic role assignment. Depicted-event information helped children to get rid of their initial preference for the preferred SVO structure when interpreting OVS sentences. However, visual context effects were subtly delayed in children (vs. adults), and varied as a function of their accuracy and cognitive capacity.

Keywords: eye tracking; child language comprehension; visual context; depicted events.

Introduction

How visual context affects adult language comprehension has been extensively investigated. Adults can rapidly use visual context information such as referential contrast (e.g., Tanenhaus, Spivey, Eberhard, & Sedivy, 1995) and depicted events (Knoeferle, Crocker, Scheepers & Pickering, 2005; Knoeferle, Habets, Crocker, & Münte, 2008) for syntactic structuring and disambiguation. By contrast, for children's use of visual context in (real-time) language comprehension, there are conflicting results (Nation, Marshall & Altmann, 2003; Sekerina, Stromswold & Hestvik, 2004; Snedeker & Trueswell, 2004; Trueswell, Sekerina, Hill & Logrip, 1999; Weighall & Altmann, 2010).

In relating language to visual context, adults' adhere to what's been dubbed the 'referential principle' of language processing: When more than one syntactic analysis is offered in parallel, the referentially supported (vs. unsupported) analysis is favored. Evidence for this claim comes from a study by Tanenhaus et al. (1995). Participants heard a sentence such as *put the apple on the towel into the box* and inspected related objects (e.g., an apple on a towel, another apple on a napkin, an empty towel, and a box). The prepositional phrase *on the towel* could temporarily either modify *the apple* or (ultimately incorrect) attach to the verb phrase *put*. When two apples (on a towel vs. on a napkin) were present, the phrase *on the towel* permitted adults to rapidly identify which apple the sentence was about, and to avoid a temporary misinterpretation of *on the towel* as the action destination. This was evidenced through eye gaze: adults rapidly inspected the apple on the towel and never the empty towel destination, much like they did for structurally unambiguous baseline sentences (*put the apple that's on the towel into the box*). They thus favored the referentially supported modifier over the referentially unsupported destination interpretation (Tanenhaus et al., 1995).

By contrast, 5-year old children (4;8 to 5;1 years) couldn't use the referential visual context and the underlying referential principle for online parsing decisions when tested with similar visual contexts and sentence ambiguity. In a study by Trueswell et al. (1999), participants either heard a locally structurally ambiguous sentence such as *Put the frog on the napkin in the box* or an unambiguous sentence such as *Put the frog that's on the napkin in the box*. For the ambiguous sentence, the prepositional phrase *on the napkin* can either modify the noun (location) or attach to the verb and specify the destination of the action (destination). Children inspected a table with either only one possible referent for *frog* in the 1-referent condition (e.g. a frog on the napkin, an empty napkin, a distractor and a box) or with two referents in the 2-referent condition (e.g. a frog on the napkin, another frog, an empty napkin and a box).

When 5-year old preschoolers heard the prepositional phrase *on the napkin* in the ambiguous instructions they frequently looked at the incorrect destination (the empty napkin) in both one-referent and two-referent contexts. This was interpreted as indicating that they – unlike adults in Tanenhaus et al. (1995) – incorrectly interpreted this phrase as the destination for *put*. It suggested children did not utilize the Referential Principle to guide their interpretation of the first noun phrase and the ensuing prepositional phrase. Moreover, their actions indicated that they never revised this initial misanalysis: On 60% of the trials the children performed an action that involved the incorrect destination

(e.g., moving a frog to the empty napkin before putting it in the box). The difficulty was attributable to ambiguity rather than structural complexity, since children mostly performed the action correctly (i.e., they moved the frog into the box instead of moving it to the empty napkin) with unambiguous sentences in a 2-referent context if they had moved the target animal first (the frog on the napkin).

The authors attributed the lack of visual context effects to a general inability of children to revise initial commitments made during syntactic structuring and semantic interpretation. Because five-year-old children have more limited processing capacities than adults, it may be difficult for children to make use of the referential principle (in real time) for resolving temporary syntactic ambiguities.

However, Meroni and Crain (2011) found a referential visual context effect in a post-sentence act-out task with a subtly-adjusted experimental design and procedure. Children inspected a two-referent context in which two frogs were placed on different-colored napkins. They were asked to close their eyes while listening to spoken sentences such as *Put the frog on the red napkin into the box*. In this setting, 3-to 5-year-olds exhibited adult-like performance in a post-sentence act-out task (i.e., they moved the frog that was on the red napkin directly into the box). The authors concluded that the changes in experimental setup and procedure enabled young children to inhibit their incorrect syntactic commitment and semantic interpretation. However, they only measured children's performance in a post-sentence task, and not the moment-by-moment processes of children's online language comprehension.

We argue that another reason for the lack of visual context effects in the study by Trueswell et al. may be potential difficulties in using visual referential contrast. Children heard *Put the frog on the napkin into the box*. At *the napkin* they had a choice in referential processing between looking at the empty napkin (a potential referent for *napkin*) or at the frog that's on the napkin. Since even young children rapidly fixate the picture of a word they have recognized (Hollich, Hirsh-Pasek & Golinkoff, 2000), it may be that children pursue primarily a "referential" strategy when hearing *on the napkin*, and prefer to look at the single empty napkin (vs. another object on a napkin). If children employ such a strategy, it would garden-path them even more since it would direct their gaze to the incorrect destination (the empty napkin). Thus, to the extent that mapping words onto objects (i.e., referential processes) governs child language comprehension, we cannot exclude that the way in which words in the utterance related to objects in visual context, rendered the use of information from that visual context unnecessarily difficult for children.

Related studies have examined the effects of other kinds of information in visual context (depicted action events) on children's comprehension accuracy for difficult-to-understand relative clause sentences (Weighall & Altmann, 2011). Children heard either a center-embedded (*The cat that bumped the bear will hug the cow*.) or a right-branching (*The cow will hug the cat that bumped the bear*.) relative clause sentence. They saw clipart pictures either of the actions described by the relative clause (e.g., an orange cat bumping a bear and a striped cat bumping a sheep) or of several referents but without the actions (e.g. a bear, a sheep, an orange cat and a striped cat). When events were (vs. were not) depicted in visual context, children's accuracy on post-sentence comprehension questions was higher. These finding suggest that visual context can improve children's comprehension of difficult relative clauses and demonstrated that children can utilize extra-linguistic information such as depicted events at least for a post-sentence comprehension task.

Overall and specifically for *real-time* syntactic structuring and thematic role assignment, however, we still know little about how children use events depicting who-does-what-to-whom. We thus examined the effects of depicted events on children's versus adults' thematic role assignment in structurally unambiguous SVO and OVS sentences. Five-year-old children had difficulty in understanding unambiguous non-canonical German OVS sentences (Dittmar, Abbot-Smith, Lieven & Tomasello, 2008), and the same is true for adults (e.g., Matzke, Mai, Nager, Rüsseler, & Münte, 2002). For adults, we already know that for understanding structurally ambiguous or unambiguous OVS sentences, depicted action events can rapidly guide their visual attention and inform thematic role assignment (i.e., shortly after the verb that mediates relevant events, see Knoeferle et al., 2005, Knoeferle, 2007).

Examining effects of depicted events in children is interesting since we don't yet know whether events can rapidly guide their visual attention and thematic role assignment. Furthermore, action events involve a different relationship between the utterance (e.g., a verb such as "pushes") and visual context (e.g., a depicted action and its associated agent) than referential contrast. When children hear the verb – and assuming they incrementally establish reference to objects as has been found – they should rapidly look at the matching action, and notice its associated agent, which could help them to disambiguate / interpret the utterance.

If the lack of a visual context effects in young children results from the complexity of referential contrast rather than children's inability to use non-linguistic visual context, then we should find effects of visual context on thematic role assignment when the content of that visual context has a more direct relationship with words in the utterance (e.g., a verb that can be associated directly with a matching action). Furthermore, the studies by Meroni and Crain (2011) and Weighall and Altmann (2011) only reported children's performance in act-out and comprehension tasks, but not their continuous visual attention during spoken sentence comprehension. Continuous measures such as eye tracking can reveal more about language comprehension in real time. This measure provides a means for examining the moment-by-moment processes of children's spoken language comprehension, in the relatively natural situation of acting out spoken instructions, or answering questions.

To assess visual context effects on children's (vs. adults') processing of German SVO and OVS sentences in depicted event (vs. no-event) contexts, we conducted two eye-tracking studies - one study with children (Exp1) and the other with adults as participants (Exp2). We expected that if children benefit from the depicted events, they would show an adult-like gaze pattern and a higher accuracy in the comprehension task when events are present (vs. absent), particularly for the difficult OVS sentences. In addition, we obtained response accuracy in a post-sentence comprehension task, to ground our interpretation of eye-movement pattern in ultimate comprehension success. We also conducted cognitive tests to get insight into potential variation of context effect as a function of children's cognitive resources. We expected children with higher WM scores would show more adult-like gaze pattern.

Experiments 1 and 2

Participants

Thirty-two kindergarten children (15 4-year olds and 17 5-year olds, range: 4-5;10) took part in Experiment 1 and received a small toy for their participation. Thirty-two students of Bielefeld University received four euro each for taking part in Experiment 2. All participants had German as their only mother tongue and normal or corrected-to-normal vision. All were unaware of the experiment purpose. Children, one of their parents (Experiment 1), and adult participants (Experiment 2) all gave informed consent.

Materials

We created 16 experimental items from 64 clipart pictures edited with commercially available graphics programs, and 64 sentences. An item consisted of four images and sentences. There were two version of an image; one depicted three characters as performing actions ('event', Pictures 1a and 1c, Table 1); the other depicted the same three characters without actions ('noevent', Pictures 1b and 1d). Each image was presented together with either a structurally unambiguous spoken subject-verb-object (Table 1, (1a) and (1a') SVO) or object-verb-subject sentence ((Table 1, (1b) and (1b') OVS)). In recording the sentences we took care to use neutral prosody. The design thus included two within-subject factors: *case marking* (SVO vs. OVS) and *event depiction* (event vs. no-event).

We counterbalanced the event roles of the characters; depiction of actions did not change. For example, one picture (Picture 1a) depicted the bear as pushing the bull, and the worm as painting the bear. The bear in the middle was thus role-ambiguous (agent and patient); the bull was the patient of the pushing action, and the worm the agent of the painting action. On the counterbalancing picture (Picture 1c), the bear was still role-ambiguous but now the worm was the patient of the painting action, and the bull the agent of the pushing action. This ensured visual characteristics of the characters could not confound looks to them as role fillers in an event. The two picture versions were each presented with a SVO and an OVS sentence. Event versus no-event pictures were presented with the same sentences (see e.g., Pictures 1a and 1b in Table 1).

Table 1: Example materials

Picture	Condition	Sentence
1a 	SVO-event	(1a) *Der Bär schubst sogleich den Stier.* The bear (subj) pushes immediately the bull (obj). 'The bear pushes immediately the bull.'
	OVS-event	(1b) *Den Bär malt sogleich der Wurm.* The bear (obj) paints immediately the worm (subj). 'The bear is immediately painted by the worm.'
1b 	SVO-noevent	(1a) *Der Bär schubst sogleich den Stier.* The bear (subj) pushes immediately the bull (obj). 'The bear pushes immediately the bull.'
	OVS-noevent	(1b) Den Bär malt sogleich der Wurm. The bear (obj) paints immediately the worm (subj). 'The bear is immediately painted by the worm.'
1c 	SVO-event	(1a') *Der Bär malt sogleich den Wurm.* The bear (subj) paints immediately the worm (obj). 'The bear paints immediately the worm.'
	OVS-event	(1b') *Den Bär schubst sogleich der Stier.* The bear (obj) pushes immediately the bull (subj). 'The bear is immediately pushed by the bull.'
1d 	SVO-noevent	(1a') *Der Bär malt sogleich den Wurm.* The bear (subj) paints immediately the worm (obj). 'The bear paints immediately the worm.'
	OVS-noevent	(1b') *Den Bär schubst sogleich der Stier.* The bear (obj) pushes immediately the bull (subj). 'The bear is immediately pushed by the bull.'

We also counterbalanced character orientation to ensure that the role-ambiguous middle character was oriented equally often to the left as to the right for the experiment items. In addition to the 24 experimental items, we constructed 8 filler items. Each filler item had a scene with two characters and a sentence accompanying it. Two started with an adverbial phrase; two with an unambiguously object case-marked noun phrase; two started with a subject case-marked noun phrase; two had two characters doing the same action. Sixteen additional fillers were created for the adults in Experiment 2. Eight of them had the same structure as the fillers in Experiment 1. Of the remaining 8, 4 had pictures that showed three characters doing different actions and 4 depicted only one character. The fillers ensured that the scene did not always depict three characters and that the verb was not always in the second position.

From the four conditions, and their two counterbalancing versions (for character role and orientation, see above), sixteen experimental lists were created, each consisting of 16 experiment and 8 filler items in Experiment 1 (children) and 16 experiment and 24 filler items in Experiment 2 (adults). Apart from the filler trials and the instructions (see Procedure), there were no differences between the child and adult experiments. Each participant saw only one of the four conditions of each item and the same number of items in each condition. Item order was pseudo-randomized individually for every participant. Adults saw at least one filler item in between two experimental trials.

Procedure

An EyeLink1000 remote eye-tracker with a sampling rate of 500 Hz monitored participants' eye movements. Images were presented on a 22" LCD color monitor at a resolution of 1680×1050 pixels concurrently with the spoken sentences. We only tracked the right eye, but viewing was binocular. In Experiment 1, each child was instructed to play a game. In this game, children were asked to inspect the images and to listen to the sentences. After each trial, they heard a question about the previous sentence and were asked to try to answer it correctly. In Experiment 2, adult participants were instructed to listen to the sentences and inspect the pictures, and to answer the question accurately. For adults, we devised a story to cover our actual goals. The cover story was that the experiment examined how well adults can understand and concentrate on child materials.

Each trial started with the display of a central fixation dot followed by a picture. After a 2000-ms picture preview time, the sentence was played via speakers. Five hundred milliseconds after sentence offset, any depicted actions were removed (if present) and participants only saw the three characters. With only the characters present, a spoken question asked for either the subject or the object of the verb in the previous sentence (for example, *Wer malt hier?/ Wer wird hier gemalt?*, 'Who paints?/ Who is painted?'). Participants answered by naming the correct character. At the start of the experiment, each participant was shown two example images and sentences. Next, participants were set up and calibrated manually using a five-point fixation stimulus in Experiment 1 and a nine-point fixation stimulus in Experiment 2. The black dot for adult calibration was replaced by a smiley to attract children's attention in Experiment 1. The EyeLink software validated calibration; if validation was poor, calibration was repeated until it was good. Between trials, participants fixated a centrally-located smiley (children) or black dot (adults). This allowed the software to perform a drift correction if necessary. The experiment lasted approximately 25 mins. After the eye-tracking part children completed a working memory (WM) test (a word ordering test; Kaufmann Assessment Battery for Children), and adults were debriefed.

Analysis

We defined three analysis time windows: the verb region (from verb onset until its offset); the adverb region (from adverb onset until its offset) and the NP2 region (from NP2 onset until its offset). We coded participants' fixations to four areas of interest in the scene: the agent (e.g., the worm in Table 1 Pictures 1a and 1b); the patient (e.g., the bull in Table 1, Pictures 1a and 1b); the role-ambiguous middle character (the bear), and the background. Of those, the agent and patient were our target areas of interest. The proportions of fixation on the target areas of interest (the patient and the agent) were entered into log-ratio analyses (c.f., Arai, van Gompel & Scheepers, 2007; Carminati, van Gompel, Scheepers, & Arai, 2008; Knoeferle, Carminati, Abashidze & Essig, 2011). We computed mean log gaze probability ratios for the agent relative to the patient *ln* (*P* (agent)/*P* (patient)) for each condition and time window. Then we entered the log probability ratios into a 2 (*case marking*) × 2 (*event depiction*) repeated measures ANOVA. Separate models were fitted for log-ratios averaged over participants and items. We report the *p*-values of these analyses.

Results Experiment 1

Response accuracy Children's overall comprehension accuracy was 67%. Their accuracy was relatively high for SVO sentences in both the event and the no-event condition (87% vs. 77%). For OVS sentences, by contrast, their accuracy was higher when events were (vs. weren't) depicted (60% vs. 42%). A 2 (*case marking*) × 2 (*event depiction*) repeated measures ANOVA revealed a significant main effect of case marking ($p < .001$) and of event depiction ($p < .001$) by subjects in the absence of a reliable interaction. Spearman's *rho* confirmed that response accuracy was not significantly correlated with age.

Eye-movement results Figures 1(a) and 1(b) plot the mean log-gaze probability ratios (participants' means) of correctly answered trials per condition at the verb and adverb regions respectively. For the verb region effects involving the independent variables were not reliable (Fig. 1(a)). By contrast, during the adverb region, children looked more at the patient entity for SVO (vs. OVS) sentences and more at the agent entity for OVS (vs. SVO) sentences, but only

when events were depicted (Fig. (1b)). When there were no depicted events, the eye-gaze data showed that children didn't inspected the correct target object until it was named at NP2 region. Inferential analyses corroborated these effects, as evidenced by an interaction between case marking and event depiction for the adverb region that was significant by subjects ($p < .005$). Thus, case marking at the first noun phrase was not sufficient for incremental thematic role assignment in children; the events, however, enabled correct thematic role assignment. Note that neither case marking nor prosody could account for the different gaze pattern observed in the event vs. no-event conditions since sentences were identical between these two factor levels.

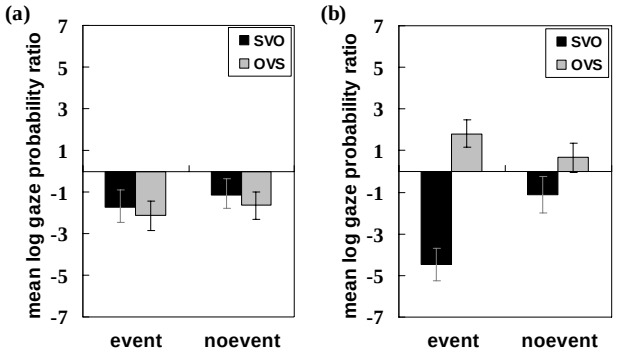

Figure 1: Children's mean log gaze probability ratios *(ln(P(*agent*)/P(*patient*))* per condition at the (a) verb region and (b) adverb region for correctly-answered trials in Experiment 1. Positive values indicate more looks to the agent; negative values indicate more looks to the patient.

We further split the subjects into high vs. low groups based on their median accuracy and working memory scores. The interaction between accuracy, case marking and events was significant ($p < .05$), and there was a non-sinificant trend towards an interaction between WM, case marking and events ($p < .15$). Separate analyses with the high vs. low groups showed that the interaction between case marking and events was significant by subjects for both the high accuracy ($p < .005$) and high working memory groups ($p < .05$) and marginal by items for the high accuracy group ($p < .1$). Response accuracy was significantly correlated with WM scores *(*Spearman's *rho, p* < 005**).** Furthermore, there was a hint that children's gaze pattern at the adverb region was correlated with their accuracy data as well (Spearman's *rho, p* < .1*)*.

Results Experiment 2

Response accuracy Adults' accuracy was high (97%). For SVO sentences, their accuracy was 100% in both the event and the no-event condition. Their accuracy for OVS sentences was also high with no reliable difference for the event versus no-event condition (96% vs. 94%). Analyses confirmed a main effect of case marking ($p < .005$), but not of event depiction and no interaction of these two factors.

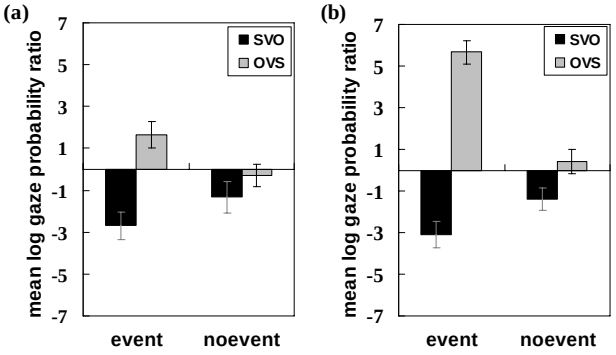

Figure 2: Adults' mean log gaze probability ratios *(ln(P(*agent*) / P(*patient*)* per condition at the (a) verb and (b) adverb region for correctly answered trials in Experiment 2. Positive values indicate more looks to the agent; negative values indicate more looks to the patient.

Eye-movement results Figures 2(a) and 2(b) present the mean log-gaze probability ratios (participants' means) for correct trials per condition at the verb and adverb region respectively for Experiment 2. Adults' (vs. children's) gaze data revealed an even earlier visual context effect, and a reliable interaction of case marking and event was confirmed at the verb (both *ps* < .01). Adults began to look more at the target object for both SVO and OVS sentences at the verb when events were (vs. weren't) depicted (Fig. 2(a)). Figure 2(b) illustrates the continued visual context effects during the adverb, confirmed by a reliable interaction of case marking and event depiction (*ps* < .001).

Discussion

We assessed how visual context information such as depicted events influences children's online language comprehension and whether children can use visual context for language comprehension to the same extent as adults. To this end we recorded both children's (Experiment 1) and adults' (Experiment 2) eye movements to characters in clipart pictures as they listened to German subject-verb-object (SVO) and object-verb-subject (OVS) sentences. The clipart pictures depicted (vs. didn't depict) who-does-what-to-whom through action events.

The results of Experiment 2 confirm the view that visual context can influence adults' visual attention and sentence processing incrementally and rapidly. Adults looked more often toward the patient character for SVO (vs. OVS) sentences and more often towards the agent character for OVS (vs. SVO) sentences when events were depicted. This gaze pattern replicated the qualitative gaze pattern and its time-course in related adult studies (e.g., effects of case marking and event depiction in Knoeferle, 2007). Not unsurprisingly, effects for the unambiguous sentences in

Experiment 2 occurred slightly earlier (at the verb) than for initially structurally ambiguous sentences (Knoeferle et al., 2005, post-verbally). Adults' response accuracy was expectedly high in both SVO and OVS sentences.

Children's accuracy for the difficult OVS sentences was higher when action events were (vs. were not) depicted (Experiment 1). Thus, action events can improve their comprehension of unambiguous non-canonical OVS sentences for which they are otherwise at chance (Experiment 1 and Dittmar et al., 2008). Unlike suggested by prior results for locally structurally ambiguous sentences (e.g. Trueswell et al., 1999; Snedeker & Trueswell, 2004) we conclude that visual context information can help children to overcome an initial structural (SVO) preference. Future research will address whether our findings extend to locally structurally ambiguous sentences.

Furthermore, effects of the events on children's visual attention and syntactic structuring emerged *incrementally*. At the post-verbal adverb and thus one region before mention of the target entity, 5-year-old children fixated the patient more often when hearing a SVO (vs. OVS) sentence and the agent when hearing an OVS (vs. SVO) sentence but only when action events were depicted. Children's gaze pattern during the adverb suggests that visual context information (depicted agent-action-patient events) can rapidly influence their sentence processing and syntactic structuring for unambiguous but non-canonical sentences.

The interaction of case marking and event in the high (but not low) WM and accuracy groups suggests that visual context effects are sensitive to children's cognitive capacities and that they may co-vary in their emergence with increases in cognitive resources. Overall, however, visual context information can inform syntactic structuring and thematic role assignment incrementally and rapidly in both 5-year old children and adults.

Acknowledgments

This research was funded by the Cognitive Interaction Technology Excellence Center (German research foundation, DFG) and a Fulbright fellowship to LK. We thank Linda Krull and Eva Mende for help with stimulus preparation and data collection; Maria Nella Carminati for advice on the analyses; and Helene Kreysa for advice on the Experiment Builder software. We also thank all the participating families and students for their support.

References

Arai, M., van Gompel, R., & Scheepers, C. (2007). Priming ditransitive structures in comprehension. *Cognitive Psychology, 54*, 218–250.

Carminati, M. N., Gompel, R. P. G. van, Scheepers, C., & Arai, M. (2008). Syntactic priming in comrpehension: the role of argument order and animacy. *JEP: LMC, 34*, 1098–1110.

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*, 1152-1167.

Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs for the Society for Research in Child Development*, 65 (3, Serial No. 262).

Knoeferle, P. 2007. "Comparing the time-course of processing initially ambiguous and unambiguous German SVO/OVS sentences in depicted events", in: R. Gompel van, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movement research: insights into mind and brain.* Oxford: Elsevier, 517 - 533

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

Knoeferle, P., Habets, B., Crocker, M. W., & Münte, T. F. (2008). Visual Scenes Trigger Immediate Syntactic Reanalysis. *Cerebral Cortex, 18*, 789-795.

Knoeferle, P., Carminati, M., Abashidze, D., & Essig, K. (2011). Preferential inspection of recent real-world events over future events. *Front. Psychology 2:376. doi: 10.3389/fpsyg.2011.00376*

Melchers, P. & Preuß, U. (2009). *Kaufman Assessment Battery for Children (deutsche Version) (8., unveränd. Aufl.).* Frankfurt/M.: Pearson Assessment.

Meroni, L. & Crain, S. (2011). "How Children Avoid Kindergarten Paths". In Edward Gibson and Neal Pearlmutter (Eds.), *The processing and acquisition of reference*. Cambridge, MA: MIT Press.

Nation, K., Marshall, C. M., & Altmann, G. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. J*ournal of Experimental Child Psychology, 86*, 314–329.

Sekerina, I., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language, 31*, 123-152.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238 –299.

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.

Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergartenpath effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89 –134.

Weighall, A.R. & Altmann, G.T.M. (2010) The role of working memory and contextual constraints in children's processing of relative clauses. *Journal of Child Language, 38*, 579-605.