The influence of depicted event scenes on written comprehension of locally ambiguous sentences

Pia Knoeferle[1], Matthew W. Crocker[1], & C. Scheepers[2]
knoeferle@coli.uni-sb.de
[1]Saarland University, [2]University of Dundee

Linguistic context has been found to influence comprehension of written sentences (Crain & Steedman, 1985; Altmann & Steedman, 1988). Similarly, visual contexts (visual referential context, depicted agent-action-patient events) have been shown to influence the auditory comprehension of locally ambiguous sentences (Tanenhaus et al., 1995, Knoeferle et al., 2003). One interpretation for the findings by Tanenhaus et al. (1995), and Knoeferle et al. (2003) is that the visual scene facilitates the comprehension of typically disfavored sentence structures, similar to the way in which a linguistic context may reduce comprehension difficulty in reading.

In the present experiment, we investigate whether the attentional eye-movement patterns observed in the visual world experiments by Knoeferle et al. (2003) reflect indeed processes of incremental thematic role-assignment. The method chosen was cross-modal comprehension where written sentences were preceded by *visual* contexts. The aim was to clarify how cross-modal comprehension proceeds at the visual-linguistic interface online, and to confirm our interpretation of what eye-movements in visual scenes reveal about comprehension processes: Are they truly indicative of language comprehension processes? In order to investigate this issue, we monitored eye-movements while people were reading initially ambiguous English main verb (MV)/reduced relative (RR) sentences after they had inspected a visual scene. The scene that preceded the sentence either showed the events described by the respective sentence, or did not display the events. There were hence four conditions, crossing the factors "Visual Context Type"(depicted event/no depicted event) with "Sentence structure" (MV/RR) (see Examples (1a), (1b), (2a), and (2b)). The Depicted Event images (1) always showed two events, one described by the MV sentence (ballerina-splashing-cellist), the other corresponding to the RR clause (fencer-sketching-ballerina). When the scene showed such events, thematic role-relations between the event participants were available from the scene prior to reading the sentence. If the depicted event scenes do indeed reduce comprehension difficulty for the RR clauses as compared to MV structures in the Depicted-Event conditions (1) versus the No-Depicted-Event conditions (2), then this should manifest itself in reduced reading times on the second noun phrase/by-phrase of the RR clauses in (1b) versus (2b).

An analysis of data from 32 participants confirmed that this prediction was borne out. We found a significant interaction of "Visual Context Type" and "Sentence Structure" in Regression Path Duration on the NP2/by-phrase region. Planned comparisons revealed that the difficulty associated with RR structures was smaller in the "Depicted Event" condition than in the "No Depicted Event" condition.

In addition, we found a significant interaction of "Visual Context Type" and "Sentence Structure" on the post-verbal region (*apparently*) with the opposite pattern. Reading times for this region were longer for RR clauses in (1b) than in (2b). We interpret this finding as indicative of an earlier structural revision for RR clauses in (1). This finding further confirms that the visual context reduced the processing difficulty associated with RR clauses on the prepositional phrase.

The results support the hypothesis that depicted agent-action patient events, and visual environments in general, influence both spoken and written sentence comprehension. In particular, they provide clear support for an interpretation of findings by Knoferle et al. (2003) in terms of thematic role-assignment processes. Furthermore, they indicate that results from studies which monitor anticipatory eye-movements in visual scenes do indeed reflect processes that are not unrelated to those revealed by reading-time studies.

Examples

	Visual Context Type	Sentence Structure	
(1a)	Depicted Event	MV	The ballerina splashed apparently the cellist in the white shirt.
(1b)	Depicted Event	RR	The ballerina sketched apparently by the fencer splashed the cellist.
(2a)	No Depicted Event	MV	The ballerina splashed apparently the cellist in the white shirt.
(2b)	No Depicted Event	RR	The ballerina sketched apparently by the fencer splashed the cellist.

References

Altmann, G.T.M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191-238.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological syntax processor. In: D. Dowty, L. Karttunen, & A. Zwicky (eds), *Natural Language Parsing*, (pp.320-358).Cambridge: CUP.

Knoeferle, P., Crocker, M.W., Scheepers, C. & Pickering, M. (2003). Actions and roles: using depicted events for disambiguation and reinterpretation in German and English. In: *Proceedings of the 25th Annual Conference of the Cogsci Conference*, Boston, MA.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K., & Sedivy, J.C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.