Scene-Sentence Integration:
Incremental Effects of Mismatch and Scene Complexity

Pia Knoeferle
Dept. of Computational Linguistics
Saarland University, Germany
`knoeferle@coli.uni-sb.de`

Monica Rodriguez
Dept. of Computational Linguistics
Saarland University, Germany
`monic@coli.uni-sb.de`

Abstract

We monitored eye movements in a scene during spoken sentence comprehension to investigate the effects of different types of scene-sentence mismatch (action vs. role relations) and of scene complexity on comprehension. Gaze analyses revealed rapid effects of both role relations mismatch and scene complexity, while effects of action mismatch were slightly delayed.

1 Introduction

Verification-task studies have reported longer response latencies (e.g., Just & Carpenter, 1971) and gaze durations (Underwood, Jebbett, & Roberts, 2004) for resolution of a sentence-picture mismatch compared with a match, suggesting a mismatch is more complex to process than a match. We extended the mismatch approach by investigating how different types of scene-sentence mismatch (action versus role relations mismatch, Experiment 1), as well as scene complexity (Experiment 2) affect incremental thematic interpretation. To obtain further insights into the time-course of scene-sentence integration, we monitored participants' eye movements in a scene during comprehension of a related utterance.

2 Experiment 1

2.1 Method

Twenty-four German native speakers with normal vision received each five euro for experiment participation. There were 24 items. Presenting the sentence in Table 1 with the four images in Fig. 1 (A to D) created four conditions (see Table 1).

For counter-balancing reasons, one item had two sentences and four images, resulting in eight
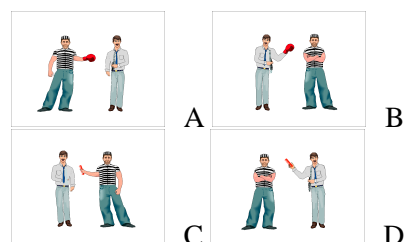

Figure 1: Example Item Images

Sentence & Fig.	Role	Action
1A Der Sträfling boxt gerade den Flötisten	Match	Match
'The convict (S) punches currently the flautist (O)'		
1B Der Sträfling boxt gerade den Flötisten	Mism.	Match
'The convict (S) punches currently the flautist (O)'		
1C Der Sträfling boxt gerade den Flötisten	Match.	Mism.
'The convict (S) punches currently the flautist (O)'		
1D Der Sträfling boxt gerade den Flötisten	Mism.	Mism.
'The convict (S) punches currently the flautist (O)'		

Table 1: Example Item Sentences

experimental lists. Items were rotated across lists such that no participant saw more than one version of each item, and such that each condition appeared equally often in each list. Consecutive experiment trials were separated by at least one of 48 filler trials. An SMI EyeLink I head-mounted tracker monitored participants' gaze in the scene during spoken comprehension. There was no verification task. Rather, participants were instructed to try to understand both sentences and depicted scenes. For half of the 48 filler trials, a written yes/no question about the sentence ensured that people performed a comprehension task. We report analyses of gaze durations that started in the ADV (from adverb onset to the onset of the second noun phrase), and NP2 regions. During these time regions the available scene and utterance information should permit resolution of both the action and role mismatch. If these two types of mismatch rapidly affect thematic interpretation, then their effects should be reflected in the inspection

durations on the target characters (the scene agent, 'the convict', and patient, 'the flautist') during the analyses regions.

2.2 Results and Discussion

The key finding is the rapid effect of the role relations mismatch on thematic interpretation as evidenced by an interaction between target character (agent, patient) and role mismatch in the ADV region ($ps < 0.01$, see Fig. 2). People inspected the patient longer than the agent for a role match (C1 & C2, Fig. 2), while there was no such difference for a role mismatch. In contrast, there was no reliable effect of action mismatch in the ADV region. For the NP2 region, there were no reliable effects of the mismatch regarding gaze durations on the target characters.
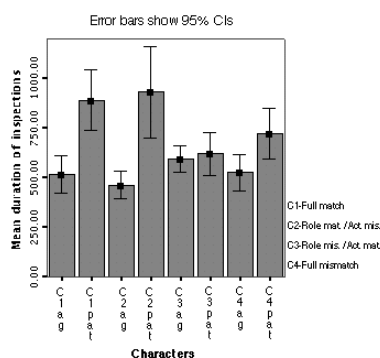

Figure 2: Mean inspection durations to the target characters for the ADV region in Experiment 1

3 Experiment 2

Experiment 2 reused the materials from Experiment 1 but retained only the action mismatch to verify its effects independent of the role relations mismatch. We further examined the influence of scene-complexity (simple vs. complex) on scene-sentence integration. Simple scenes contained the two target characters (agent, patient) of Experiment 1 and four distractor objects. Complex scenes showed an additional three characters.

3.1 Method

Thirty-two further participants from the same population as in Experiment 1 were each paid five euro. Procedure, task, and the analyses regions were the same as in Experiment 1. In addition, we examined early effects of scene complexity by analyzing the duration of inspections that started after NP1 and before verb onset.

3.2 Results and Discussion

There was a main effect of scene complexity for NP1 ($ps < 0.01$), with longer inspection durations on target characters (agent, patient) for simple than complex images. During the ADV region we found no effects of either action mismatch or scene complexity. For NP2, there was an interaction of mismatch and target character ($ps < 0.001$): people fixated the patient longer than the agent for the action-match conditions (C1 & C3). For action-mismatch conditions (C2 & C4), in contrast, inspection duration on the agent and patient did not differ (Fig. 3).
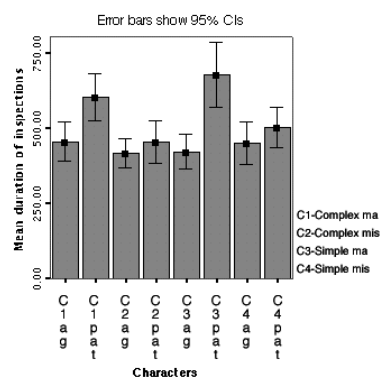

Figure 3: Mean inspection durations to target characters for the NP2 region in Experiment 2

4 Conclusions

Taken together, our findings support the view that scene-sentence integration takes place incrementally. There were, however, differences in the time course of processing actions and role relations mismatch: While the role relations mismatch influenced thematic interpretation post-verbally, effects of the action mismatch only affected thematic interpretation later, during the NP2 region. Scene complexity did not interact with action mismatch, but influenced the inspection duration of the target characters during NP1.

References

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

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