Eye-tracking evidence for online processes in the comprehension of referential expressions

Anke Karabanov, Peter König, Peter Bosch University of Osnabrück, Institute of Cognitive Science {akaraban@uos.de, pkoenig@uos.de, pbosch@uos.de}

Noun phrase reference (and co-reference) plays a central role in theoretical linguistics, standing as it does at the interface between syntax, semantics, and pragmatics. Important questions concerning the cognitive processes underlying reference phenomena have, until recently, not lent themselves to direct investigation, however. In this paper we present a preliminary study, in which visual attention is used to investigate the time course of reference resolution in discourse. Psycholinguistic studies of NP reference have almost exclusively relied on indirect evidence, such a reading times, reaction times, or eye-movement during reading (cf. surveys in Nicol and Swinney 2002, Rayner and Clifton 2002, Garnham 1999), detouring attention from central questions of interest. Experimental methods recently developed mainly by Tanenhaus and his colleagues (Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy, 1995) under the label "Visual World Paradigm" provide for more direct investigation of the time course of language processing. This paradigm has been used to investigate the processing of lexical information (Tanenhaus, Magnuson, Dahan, and Chambers, 2000), gender information (Arnold, Eisenband, Brown-Schmidt, and Trueswell, 2000) and sentence-internal co-reference (Runner, Sussman, and Tanenhaus, 2003). In the experiment reported here, we make use of an expanded visual world setup to investigate questions concerning comprehension processes underlying NP reference. In particular we are concerned with the following: Are definite full NPs understood faster than pronouns? Do pronouns require additional time for "resolution" and is this related to the "distance" between a pronoun and its antecedent or the specific nature of the syntactic or semantic relationship between pronoun and antecedent? Or are pronouns interpreted "immediately", just like proper names, as was argued already by Tyler and Marslen-Wilson (1980)?

Eye movements were recorded with a head-mounted eye-tracking system while subjects were viewing pictures on a monitor and listening to short pieces of pre-recorded narrative discourse. Each discourse consisted of three sentences, the first of which introduced the whole scene without referring to any object directly; the second sentence introduced two referents explicitly, each with a definite full noun phrase; and the third sentence picked up the two referents already referred to, this time via pronouns, and introduced a third referent with a full definite NP, as in example (1). Heute ist Markt im Dorf. Die Marktfrau streitet sich mit dem Arbeiter. Sie sagt jetzt gerade, daβ er ihr nun das neue Fahrrad zurückgeben soll, das er sich geliehen hat.

'It's market day in the village. The market woman is having an argument with the worker. She's saying that he should give her the new bike back that he borrowed.'

Ten adult subjects viewed 10 different scenes and listened each to 20 sets of sentences plus 30 sets of filler materials. In contrast with earlier Visual World studies we did not use cartoons or line drawings but photographs of Playmobil® scenes that did not just contain the three reference objects but more complete and thus more realistic settings with many additional objects not referred to in the discourse (cf. Fig.1). In contrast to Tanenhaus and his colleagues the scenes contained a large number of objects, and subjects could freely visually explore the scenes. Results show that full noun phrases and anaphoric pronouns are immediately followed by an increase in the subject's fixations on the corresponding referent in the visual scene (Fig.2). These fixations reach their peak at about 1500 ms after the onset of the referential expression, both for pronouns and for full NPs. However, the peak in fixations caused by the full NPs (mean value 42 %) was significantly higher than the fixation peak caused by the anaphoric pronouns (mean value 29%). Also there was more variation in the number of fixations caused by the different full NPs (between 35% and 49%) than between the number of fixations caused by anaphoric pronouns (between 28% and 29%). Finally, and of particular interest to linguistic theory, we observed a difference between regular anaphoric pronouns and syntactically bound personal pronouns (co-indexed and c-commanded): The latter were not followed by a significant increase in focusing and thus are apparently not interpreted referentially in the same sense as the definite full NPs and anaphoric pronouns in our experiment. Fig.3 shows the mean values of fixations (roughly) aligned to the discourse in one of the conditions. The fact that the fixation peak is reached within the same time for full NPs and anaphoric pronouns suggests that anaphoric pronouns are referentially interpreted very much like definite full NPs and that no extra processing time is needed to resolve the anaphoric reference. Together with the fact that no difference was detected between the comprehension of anaphoric pronouns closely following their antecedents and those following much later, this supports the notion that pronouns are interpreted directly with respect to referents (in a mental model, discourse representation, or here a visual model), rather than by reference to any antecedents. - Further detailed study of different types of pronouns and of referents of different saliency status, including the resolution of referential ambiguities, is clearly required to investigate the limitations of these conclusions.



Fig. 1



Fig. 2



Fig. 3

References

- Allopenna, P., J. Magnuson, and M. Tanenhaus (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. Journal of Memory and Language, 38: 419-439.
- Arnold, J., J. Eisenband, S. Brown-Schmidt, and J. Trueswell (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eye tracking. Cognition, 76: B13–B26.
- Garnham, A. (1999). Reference and anaphora. In: S. Garrod & M. J. Pickering, eds., Language processing 1999, pp. 335-362.
- Marslen-Wilson, W., and L. Tyler (1987). Against modularity. In: J.Garfield ed., Modularity in Knowledge Representation and Natural Language Understanding 1987, pp. 37-62.
- Nicol, J. and D. Swinney (2002) The psycholinguistics of anaphora. In: A. Barss, ed., Anaphora 2002, pp. 72-104.
- Rayner, K., and C. Clifton (2002). Language processing. In: D. Medin (Volume Editor) Stevens Handbook of Experimental Psychology, Volume 2, Memory and Cognitive Processes 2002, pp. 261-316.
- Runner, J., R. Sussman, M. Tanenhaus (2003). Assignment of reference to reflexives and pronouns in picture noun phrases: Evidence from eye movements. Cognition, 81.1: B1-B13.
- Tanenhaus, M., M. Spivey-Knowlton, M., K. Eberhard, and J. Sedivy (1995). Integration of visual and linguistic information in spoken language comprehension. Science, 268: 1632-1634.
- Tanenhaus, M., J. Magnuson, D. Dahan, and C. Chambers (2000). Eye movements and lexical access in spoken language comprehension: Evaluating a linking Hypothesis between fixations and linguistic processing. Journal of Psycholinguistic Research, 29: 557-580.
- Tyler, L., and W. Marslen-Wilson (1982): Processing utterances in discourse context: Online resolution of anaphors. Journal of Semantics, 1: 297-314.