What does eye-tracking reveal about children's knowledge of linguistic structure?

Petra Hendriks

University of Groningen

1. Introduction

Since the rise of generative grammar, a respectable way of investigating the nature and architecture of grammar is by studying children's knowledge of grammar. Although the analysis of children's spontaneous speech can give us an indication of children's linguistic knowledge, such studies generally fail to reveal the subtle distinctions involved in their comprehension of linguistic structure. Consequently, many acquisition studies investigating children's linguistic knowledge in sentence comprehension use experimental methods. Using standard adult experimental procedures such as reaction time measurements, reading studies, or linguistic judgment methods with children, however, has proven to be difficult and has yielded results that can be hard to interpret (Trueswell & Gleitman, 2007). Therefore, several child language researchers have begun to use eye-tracking methods, starting with Trueswell, Sekerina, Hill, and Logrip (1999). The results of this study and later eye-tracking studies suggest that recording children's eye gaze patterns while they are listening to spoken utterances in the presence of a relevant reference world may be a promising approach to studying the acquisition of sentence meaning. However, two issues should be kept in mind when performing and interpreting such eye-tracking studies. First, eye movement data associated with spoken utterances should be interpreted in the light of linguistic theory. Second, the evidence obtained by eye-tracking should ideally converge with evidence

obtained by other methods of investigation. The importance of these two issues will be illustrated on the basis of a widely cited eye-tracking study on children's knowledge of pronouns, namely Sekerina, Stromswold, and Hestvik (2004). If these two issues are ignored, incorrect conclusions may be drawn from children's gaze patterns.

2. Pronoun interpretation and linguistic theory

An extensively studied area of language acquisition is children's acquisition of object pronouns. Many comprehension studies have established that children who correctly interpret reflexives from the age of 4 or 5 on have trouble interpreting object pronouns correctly until the age of 6 or even later (e.g., Chien & Wexler, 1990, for English; Deutsch, Koster, & Koster, 1986, for Dutch). In contrast to adults, children allow for an interpretation of (1) according to which the object pronoun *him* co-refers with the local subject *the boy*:

(1) The boy_i has washed himself_i / $\text{him}_{j/*i}$

This so-called Delay of Principle B Effect has been observed on the basis of a variety of off-line experimental methods: sentence verification tasks, truth-value judgment tasks, picture selection tasks and act out tasks. This suggests that the Delay of Principle B Effect is a robust effect, at least in languages such as English and Dutch. The Delay of Principle B Effect has inspired several revisions of the original Binding Theory, illustrating the close relation between linguistic theory and language acquisition research.

Reflexives and pronouns are generally in complementary distribution. This is reflected by the original formulation of Principle A ("A reflexive is bound in its governing category")

and Principle B ("A pronoun is free in its governing category") of Binding Theory. However, there are a number of environments where reflexives and pronouns can both occur, with the same meaning. An example is (2), which allows for the reflexive as well as the pronoun inside the locative PP to co-refer with the subject of the sentence:

(2) The boy_i has placed the box behind $himself_i$ / $him_{j/i}$

The pronoun in (2) is ambiguous, and can co-refer with the sentence-internal antecedent (i.e., the subject of the sentence), or an unmentioned sentence-external antecedent available in the discourse. The reflexive in (2) must co-refer with the subject.

Reinhart and Reuland (1993) explain several well-known exceptions to the complementary distribution of reflexives and pronouns, such as (2), by reformulating the binding principles as conditions on reflexivity and reflexive-marking, and restricting the application of Principle A to syntactic predicates and the application of Principle B to semantic predicates. An alternative explanation is the optimality theoretic explanation of Hendriks and Spenader (2005/6), who propose that a hierarchy of constraints referring to binding domains of different sizes, such as locative PPs and clauses with a subject, interacts with a hierarchy of constraints pertaining to the referential content of referring expressions. Both theoretical accounts are able to explain the general pattern of distribution and interpretation of pronouns and reflexives illustrated in (1), as well as exceptions to this pattern exemplified in (2). In sentence (1), the pronoun must be disambiguated to give rise to the adult non-coreferential interpretation. Disambiguation of the pronoun by means of resource-sensitive operations such as reference-set computation (Reinhart, 2006) or bidirectional optimization (Hendriks & Spenader, 2005/6) is argued to go beyond children's limited cognitive capacities, resulting in the Delay of Principle B Effect. Because the pronoun in (2)

is ambiguous for adults, there is no need for disambiguation. Hence, neither account predicts any comprehension difficulty for children in this case.

Unfortunately, experimental investigations of children's knowledge of Principle B have mainly focused on sentences such as (1), and have rarely taken into consideration sentences like (2). Consequently, not much is known about children's interpretation of pronouns in locative PPs.

3. Using eye-tracking to assess children's knowledge

A notable exception to the general focus of acquisition research on Principle B environments is Sekerina et al. (2004), who carried out an eye-tracking study to investigate adults' and children's knowledge and processing of sentences such as (2). The participants in their eye-tracking experiment saw two pictures, and were then asked a question about the pictures. For example, one picture would show a boy and a man with a box behind the boy, and the other picture would show the same boy and the same man with the box behind the man. Participants would then hear the question "Which picture shows that the boy has placed the box behind him?". In this forced-choice picture selection task, the adult participants had to press a button as fast as they could to indicate their choice, while their reaction times were measured. The 4- to 7-year old child participants had to choose a picture by pointing to it with their finger, without any measurement of their reaction times. Therefore, instead of a question, they heard the instruction "Now point to the picture where the boy has placed the box behind him".

For adults, the obtained data from responses, reaction times and eye movements all three indicated that the pronoun was referentially ambiguous to them. Adults frequently chose the sentence-external referent, although they did have a strong preference for the sentence-

internal referent. For children, the off-line responses and on-line eye movements gave different results. Initially, the children mainly looked at the picture with the sentence-internal referent. Only after about 1000 milliseconds they started looking at the picture with the sentence-external referent. This suggests that, although the children take much longer than the adults to become aware of the ambiguity of the pronoun, they did notice this ambiguity. However, children's awareness of the ambiguity did not affect their choice in the picture selection task. In the picture selection task, children overwhelmingly chose the sentence-internal referent.

4. Children's knowledge of Principle B

The on-line and off-line performance of the adults in their study is interpreted by Sekerina et al. (2004) as providing support for a performance-based explanation, as opposed to a grammar-based explanation. According to a performance-based explanation, the interpretation of pronouns in sentences such as (2) is determined by discourse considerations. That is, these pronouns refer to the most salient, most recent, most prominent referent in the discourse, which is the subject of the sentence in Sekerina et al.'s study. However, the performance of the adults in their study is also compatible with the theoretical accounts of Reinhart and Reuland (1993) and Hendriks and Spenader (2005/6). Recall that both theoretical accounts predict pronouns in the syntactic environment in (2) to be ambiguous. If the grammar allows for a sentence-internal as well as a sentence-external referent, discourse considerations may make one of the two interpretations the preferred interpretation. As picture selection tasks do not test whether a particular interpretation is possible or not, but rather which of a given set of interpretations is the preferred interpretation, we expect such discourse considerations to

influence participants' choice. In fact, it is unclear how a performance-based explanation alone would be able to explain the observed pattern. Without a linguistic theory that accounts for why the pronoun in (2), but not the pronoun in (1), is syntactically ambiguous, pronouns are incorrectly predicted to be interpreted similarly in (1) and (2).

Regarding children's eye movement results, Sekerina et al. argue that the similarity between children's eye movement data and adults' eye movement data supports the hypothesis that children's processing mechanisms work just like adults'. Children use referential information in exactly the same way as adults, they just need more time to access discourse information during sentence processing. Again, the generative syntactic account of Reinhart and Reuland and the optimality theoretic account of Hendriks and Spenader are able to explain these results too, albeit from the perspective of the grammar. The pronoun in (2) is syntactically ambiguous for children, just as it is for adults.

The theoretical accounts of Reinhart and Reuland (1993) and Hendriks and Spenader (2005/6) make a distinction between the syntactic environment in (1), a standard Principle B context, and the syntactic environment in (2). In the latter environment, the complementary distribution between reflexives and pronouns breaks down as a result of particular properties of the principles or constraints of the grammar. Consequently, children's adult-like pattern of eye movements with sentences such as (2) does not tell us anything about their ability to interpret pronouns correctly in Principle B contexts such as (1). Sekerina et al. (2004), however, explicitly relate children's performance on sentences such as (2) to "the pattern of difficulty well-attested in the acquisition literature" (p. 148), ignoring the different treatments these sentences receive in the theoretical literature. By doing so, they suggest that, at least at the level of eye movements, the children in their study do not show any difficulty with Principle B. This syntactically unfounded suggestion may have led others to claim that studies

using eye-tracking have shown that "children as young as 4 years of age demonstrate on-line sensitivity to syntactic binding constraints" (Love et al., 2009: 287).

5. When evidence does not converge

Whereas the children in Sekerina et al.'s study displayed adult-like eye movements during the picture selection task, they gave non-adult-like responses on the task. Adults chose the sentence-external referent in 20% of cases, but children did so in only 7% of cases. To explain this dissociation between children's on-line performance and their off-line performance, Sekerina et al. hypothesize that children's eye movements reveal implicit awareness of the referential ambiguity of the pronouns in the task that develops earlier than the explicit knowledge required in a picture selection task.

However, there may be an alternative, and perhaps more plausible, explanation of the dissociation between children's eye movements and their responses on the picture selection task. A picture selection task is a forced-choice task. In Sekerina et al.'s experiment, participants were asked to choose between two pictures and determine "which of the two pictures represented the correct answer to the experimental question" (p. 136). This task becomes problematic if the sentence is ambiguous between the readings represented in the pictures. However, this is exactly what the theoretical accounts of Reinhart and Reuland (1993) and Hendriks and Spenader (2005/6) predict to be the case for the test materials used. So the two pictures match the sentence equally well according to the grammar. However, the task instructions suggest to the participants that only one of the pictures matches the sentence and the other one does not. Thus, children's deviant responses may have been caused by difficulties arising from a conflict between the task instructions and their linguistic

knowledge. This problem arises in all experiments that apply a picture selection task to ambiguous sentence materials. Note that for materials that are ambiguous for children but not for adults, such as the pronoun sentence in (1), a picture selection task may have the opposite effect and may actually help the child to arrive at the correct interpretation. In this case, the task instructions provide the child with a crucial piece of knowledge that the child may not yet have on the basis of his or her grammar, namely that only one of the interpretations is correct. This could account for the observation that in general a picture selection task seems to be 'easier' for children than a truth-value judgment task.

So the dissociation Sekerina et al. found between children's eye movements and their off-line responses may very well be a task-related effect, caused by problems that are specific to the picture selection task. Others, who assumed this dissociation to be real, have used this dissociation to support their own finding of a discrepancy between their eye-tracking results and other experimental tasks (e.g., Höhle et al., 2009). Only if these dissociations are based on carefully designed experiments whose results are supported by additional evidence, it may be concluded that the eye-tracking methodology is more sensitive than other empirical methods. However, because of the potential problems with Sekerina et al.'s picture selection task, there is no reason yet to treat eye gaze patterns as a special source of evidence for children's knowledge that overrides all other sources of evidence.

6. Conclusion

To explain the results of their eye-tracking study, Sekerina et al. (2004) suggest that children possess knowledge of Principle B, but are not yet able to apply their knowledge in a picture selection task. However, the validity of the first part of their suggested explanation was

questioned in section 4, where it was pointed out that correct performance on the test sentences does not imply similar correct performance in Principle B contexts. The validity of the second part of the suggested explanation was questioned in section 5, where it was shown that there may have been independent problems with the picture selection task in their study. So the results of Sekerina et al.'s study do not seem to lend support to the view that eye-tracking is more sensitive to children's knowledge than commonly used off-line methods such as picture selection tasks. It is possible that eye-tracking is more sensitive than other methods, but the study by Sekerina et al. does not provide evidence for this claim. This discussion emphasizes that one should be careful about how to explain eye-tracking data, especially in the absence of converging evidence.

Acknowledgments

I would like to thank the members of the Spoken Syntax Lab and the Acquisition Lab at Stanford University for insightful discussions on some of the issues addressed in this squib. Of course, all errors are my own.

References

Chien, Yu-Chin, & Kenneth Wexler (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1, 225–295.

Deutsch, Werner, Charlotte Koster, & Jan Koster (1986). What can we learn from children's errors in understanding anaphora? *Linguistics* 24, 203-225.

Hendriks, Petra, & Jennifer Spenader (2005/6). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition* 13, 319-348.

Höhle, Barbara, Frauke Berger, Anja Müller, Michaela Schmitz, & Jürgen Weissenborn (2009). Focus particles in children's language: Production and comprehension of *auch* 'also' in German learners from 1 year to 4 years of age. *Language Acquisition* 16: 36-66.

Love, Tracy, Matthew Walenski, & David Swinney (2009). Slowed speech input has a differential impact on on-line and off-line processing in children's comprehension of pronouns. *Journal of Psycholinguistic Research* 38: 285–304.

Reinhart, Tanya (2006). *Interface strategies*. Cambridge, MA: MIT Press.

Reinhart, Tanya, & Eric Reuland (1993). Reflexivity. *Linguistic Inquiry* 24, 657-720.

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

Trueswell, John, & Lila Gleitman (2007). Learning to parse and its implications for language acquisition. In: G. Gaskell (ed.), *Oxford Handbook of Psycholinguistics*. Oxford: Oxford University Press.

Trueswell, John, Irina Sekerina, Nicole Hill, & Marian Logrip (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition* 73, 89-134.