J. Child Lang. **31** (2004), 123–152. © 2004 Cambridge University Press
DOI: 10.1017/S0305000903005890 Printed in the United Kingdom

How do adults and children process referentially ambiguous pronouns?*

IRINA A. SEKERINA

Department of Psychology, College of Staten Island of the City University of New York

KARIN STROMSWOLD

Department of Psychology and the Center for Cognitive Science, Rutgers University

AND

ARILD HESTVIK

University of Bergen, Norway, and the CUNY Graduate School and University Center

(*Received* 25 *June* 2002. *Revised* 20 *March* 2003)

ABSTRACT

In two eye-tracking experiments, we investigate adults' and children's on-line processing of referentially ambiguous English pronouns. Sixteen adults and 16 four-to-seven-year-olds listened to sentences with either an unambiguous reflexive (*himself*) or an ambiguous pronoun (*him*) and chose a picture with two characters that corresponded to those in the sentence. For adults, behavioural data, responses and reaction times indicate that pronouns are referentially ambiguous. Adults' eye movements show a competition between the looks to sentence-internal

[*] This work was partially supported by National Science Foundation grants BCS-4875168 and BCS-0042561 to the second author. We thank the Rutgers University Center for Cognitive Science for use of their eye-tracking facilities provided by the NSF grant. We thank Margaret Sudhakar and Ned Norland for assisting us with preparation and conducting the experiments, Jennifer Venditti for providing her expertise in analysing the materials acoustically, and Gary Winkel for help with the statistical analyses. We also wish to thank the students from Irwing Elementary School (Speech coordinator Ellyn Atherton) and the children from the Yellow Brick Road Preschool (Director Grace Puleo) in Highland Park, NJ, for their help in running the study. We are also grateful to Eva Fernández, Patricia Brooks, Julien Musolino, Jill Razdan, and two anonymous reviewers for very helpful commentary on earlier drafts of this article. Address for correspondence: Dr. Irina Sekerina, Department of Psychology, 4S-112, CSI/CUNY, 2800 Victory Boulevard, Staten Island, NY 10314, USA. e-mail: Sekerina@postbox.csi.cuny.edu; tel: (718) 982-3760.

and -external referents for pronouns, but not for reflexives. Children overwhelmingly prefer the sentence-internal referent in the off-line picture selection task. However, their eye movements reveal implicit awareness of referential ambiguity that develops earlier than their explicit knowledge in the picture selection task. This discrepancy between performance on a looking measure and a pointing measure in the children's processing system is explained by a general dissociation between implicit and explicit knowledge proposed in recent literature on cognitive development.

INTRODUCTION

The acquisition of the two classes of pronominals, pronouns such as *him* and *her*, and reflexives such as *himself* and *herself* in English illustrated in (1) has been discussed in numerous studies that examine the question of when children come to know the contrast between these two classes. The interpretation of pronouns can be constrained either by syntactic rules such as the Binding Theory and the A-Chain Condition (Chomsky, 1981 and thereafter), or by interpretative rules such as Rule I (Grodzinsky & Reinhart, 1993) operating at the LF-discourse interface.

(1a) The boy_1 has washed himself_1.
(1b) The boy_1 has washed $\text{him}_{2/*1}$.

The basic intuition derived from the contrast in (1a) and (1b) is that reflexives and pronouns are in complementary distribution. The linguistic literature, however, cites a number of classical exceptions to the principles of the Binding Theory where this complementarity breaks down. An example involves pronouns that take an antecedent within a local domain where Principle B should rule them out, as illustrated by (2). Assuming the local domain to be the clause containing the pronoun inside the prepositional phrase (PP) in (2b), we predict that *him* cannot refer to *the boy*. Surprisingly, the sentence is referentially ambiguous, i.e. the pronoun can be coreferential either with a sentence-external referent, a discourse referent not specified in the clause (for example, a man present in the discourse), or with the sentence-internal referent, the subject of the clause, *the boy*. Moreover, the sentence-internal referent (which should be ruled out by Principle B) is intuitively the preferred interpretation.

(2a) The boy_1 has put the box behind himself_1.
(2b) The boy_1 has put the box behind $\text{him}_{1/2}$.

Tenny (1999) refers to this phenomenon as SHORT-DISTANCE PRONOUNS (SDP) and identifies two other productive environments for SDPs in English: colloquial sentences such as *I_1'll make me_1 a sandwich*, and

representational noun phrases, *Jimmy$_1$ hates stories about him$_1$ told by his cousin*.

The linguistic literature provides several accounts of why SDPs in (2b) as an exception can refer to the sentence-internal referent (Tenny, 1999; Runner, Sussman & Tanenhaus, in press). The sentence-external interpretation, on the other hand, does not present a problem from the syntactic point of view since is not regulated by purely syntactic constraints; rather, it is constrained by the Interpretation Rule, Rule I, at the LF-Discourse, one of the heuristics for the interpretation of referential expressions in terms of preceding discourse (Garrod, 1994; Gordon & Scearce, 1995). In contrast to the linguistic accounts, current processing theories predict that the sentence-internal referent being the most recent referent and the subject of the clause should be preferred (Arnold, Eisenband, Brown-Schmidt & Trueswell, 2000). In (2b), the NP *the boy* is more accessible than the discourse referent *the man* because *the boy* is the grammatical subject of the clause and occurs in the sentence itself, rather than the surrounding context. Assigning *him* to the sentence-external referent *the man* should be difficult since the 'Subject rule' will have to be overruled: the antecedent is only introduced by the discourse and is not mentioned in the sentence. Only higher-ranked linguistic constraints should be able to promote the sentence-external referent for SDPs. For example, when the sentence-external referent *the man* is promoted by the persuasive context, its discourse prominence may make it a valid competitor against the sentence-internal referent *the boy*. Thus, the psycholinguistic expectations (the Subject rule and the accessibility factors) contradict the linguistic theory for which the sentence-internal interpretation is treated as an exception. Surprisingly, there are no experimental studies in adult psycholinguistic research that either support or refute the referential ambiguity for SDPs and test the hypothesized processing preferences.

Likewise, we know very little about children's knowledge of SDPs and how they resolve referential ambiguity. Numerous language acquisition studies have shown that children have difficulties using Rule I across languages and constructions (Chien & Wexler, 1990; Avrutin, 1994) making the sentence-external interpretation difficult to access. Moreover, as young children permit coreferential readings with counterindexed NPs, allowing *him* to take *the boy* as antecedent in simpler sentences such as (1b), one might expect children to prefer the sentence-internal referent in (2b).

One explanation proposed to account for children's difficulties with sentence-external referents relies on the hypothesis that children's processing and inferential resources are limited (Wykes, 1981; Chien & Wexler, 1990; Avrutin, 1994; Reinhart, 1999). While the properties of the adult processing system have been in the focus of experimental psycholinguistics for over 30 years, we are just beginning to discover how children's

processing mechanisms develop. The important issue is whether adults and children access the same parsing routines or not. It is notoriously difficult to study children's language comprehension on-line due to a lack of child-friendly experimental techniques (*cf.* Booth, MacWhinney & Harasaki, 2000; Hahne & Friederici, 2001; Traxler, 2002). Recently, with the advent of free head-mounted eye-tracking methodology, researchers have begun investigating the development of various aspects of children's sentence processing (Trueswell, Sekerina, Hill & Logrip, 1999). No study has yet addressed the issue of children's pronoun resolution in real time. Given this gap, we investigate how four-to-seven-year-old English-speaking children process referentially ambiguous pronouns in a moment-by-moment fashion. We compare children's and adults' referent preference data using both off-line and on-line methods to empirically test for qualitative differences between child and adult processing systems.

Factors that affect processing of pronominals in adults

The difference between pronouns and reflexives in terms of processing load has recently been addressed in a set of cross-modal priming studies by Piñango, Burkhardt, Brun & Avrutin (2001). They have demonstrated that compared to reflexives, pronouns are not inherently more complex from the processing point of view: reaction times associated with referential processing of pronouns (3) and reflexives (4) did not differ in the lexical decision task (* is the probe position).

(3a) Everyone$_1$ thinks that students like him$_1$* …
(3b) The teacher$_1$ thinks that students like him$_{1/2}$* …
(4a) The boxer$_1$ who was young defended himself$_1$* …
(4b) The boy$_1$ hid a present behind himself$_1$* …

In our Experiments 1 and 2 reported below, we capitalize on these results. If no additional processing load was found for pronouns compared with reflexives in the Piñango *et al.*'s processing study, the latter can serve as the legitimate control construction for the former. Piñango and colleagues, however, did find significantly higher reaction times for referring pronouns as in (3b) than for bound variable pronouns as in (3a), and for the 'co-argument' reflexive (4a) compared to the 'logophoric' reflexive (4b). They argued for a discourse processing complexity hypothesis: processing of (3a), where interpretation of the bound variable pronoun *him* is obtained via syntactic mechanisms alone, i.e. via binding by the quantifier *everyone*, is less costly than that of (3b). In the latter case, a possible discourse interpretation of *him* via coreference with a man not mentioned in the sentence requires the processor to access information beyond that provided in the syntax, that is, information from discourse.

A large body of literature has accumulated on factors guiding referential processing in pronoun resolution in (3b), such as accessibility, order-of-mention, and recency (Garvey, Caramazza & Yates, 1974/5; Gernsbacher & Hargreaves, 1988; Crawley, Stevenson & Kleinman, 1990; Smyth, 1994). For example, Maat & Sanders (2001) showed that the Subject rule dominates pronoun resolution in short discourse fragments like *George$_1$ hit Al$_2$. He$_{1/2}$ was really mad*. However, when the competing referent is a pronoun itself as in *George$_1$ hit him$_2$. He$_{1/2}$ was really mad*, *he* is selected as the referent for *him* less often. Maat & Sanders proposed the Pronoun rule and argued that referential ambiguity resolution is regulated by the interaction of Subject and Pronoun rules.

A recent adaptation of the eye-tracking methodology to study moment-by-moment on-line processing of spoken language (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995) allows for an investigation of referential ambiguity resolution for pronouns. Runner *et al*. (in press) conducted two eye-tracking experiments with the so-called 'picture'-NPs without possessors (*Have Ken touch a picture of him/himself*) and with possessors (*Have Ken touch Harry's picture of him/himself*). Participants listened to spoken instructions and manipulated dolls in front of a display arranged with pictures of the dolls. They argued that the eye-tracking methodology is particularly suited for testing predictions of referential ambiguity resolution: the participants' eye movements in the possessor 'picture'-NPs revealed that they were regularly considering referents incompatible with the Binding Theory for sentences with reflexives, but not for those with pronouns.

Arnold *et al*. (2000) used the head-mounted eye-tracking system to record participants' eye movements as they looked at the four types of pictures representing two male characters (Donald and Mickey) or a male and female (Donald and Minnie). Either Donald or Mickey/Minnie was carrying an umbrella. Pictures were accompanied by spoken text fragments. For example, for the picture in which Mickey carries an umbrella, the text fragment was as follows:

(5a) Donald is bringing some mail to Mickey while a violent storm is beginning.

(5b) He's carrying an umbrella, and it looks like they're both going to need it.

In the example with the gender ambiguity, Arnold and colleagues found that participants did not immediately converge on an interpretation of the correct referent since Mickey is the one who is carrying the umbrella. Instead, they looked equally at both characters for a while. However, the Subject rule was strong enough to create competition between Donald and Mickey, even in the presence of unambiguous visual context, since the

subject of the clause (Donald) was mentioned first and was thus highly accessible. Then Arnold and colleagues manipulated accessibility even further by inserting an additional clause in (5a), *He's sauntering down the hill, while a violent storm is beginning*. In this case, the increased distance between the correct referent Mickey and *he* rendered Mickey inaccessible, and left the Subject rule as the default referential ambiguity resolution strategy. As a result, participants often made mistakes in these sentences by choosing the picture with the wrong referent, Donald. Consequently, participants' eye movements revealed no looks to Mickey, providing evidence for order-of-mention and recency as sources of information that guide on-line resolution of referential ambiguity.

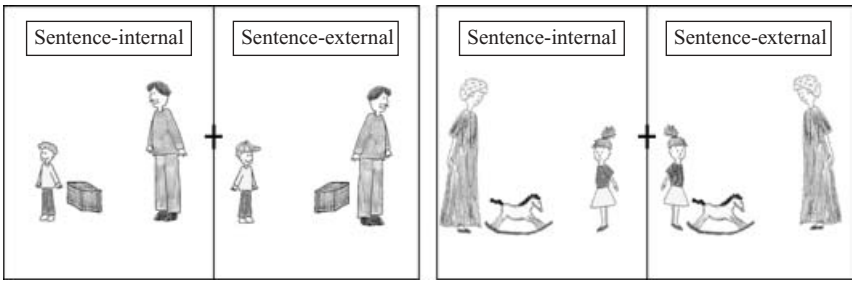
Acquisition of pronominals and development of processing mechanisms by children

From a developmental perspective, children's knowledge of the Binding Theory has been studied extensively in various languages and in many constructions (for overview, see Chien & Wexler, 1990; Grodzinsky & Reinhart, 1993). However, SDPs have not been specifically discussed. Note that from the point of view of a child acquiring English, the complex interaction of factors that allows for SDPs to be referentially ambiguous poses familiar problems of learnability (Musolino, Crain & Thornton, 2000). Theoretically, the choice of a sentence-internal vs. -external referent may be optional, obligatory, or impossible depending both on the lexical nature of the pronominal element involved (reflexive vs. pronoun) and on its position in the sentence. In the light of errors children make in violation of Principle B (1b), we would expect them to undergeneralize with SDPs: sentences containing a pronoun in the SDP context should only receive the sentence-internal interpretation. This would be in contrast to the adult grammar in which such sentences are hypothesized to allow both interpretations. The acquisition literature often ascribes this undergeneralization to the fact that preschool children have not yet mastered certain discourse mechanisms, including the Referential Principle (Crain & Steedman, 1985), Principle P (Chien & Wexler, 1990) or Rule I (Grodzinsky & Reinhart, 1993). These observations motivated the present study in which we investigated children's moment-by-moment processing of referentially ambiguous pronouns. This experiment led us to ask whether children and adults resolve this ambiguity in a similar fashion, and to explore potential qualitative differences between adults' and children's processing mechanisms.

Thornton & Wexler (1999) assume that children and adults share the same processing system, but it remains an empirical question whether that is indeed the case. Surprisingly little is known about how children process language in real-time due to the limitations in processing methodology for

studying children's language comprehension (Tyler & Marslen-Wilson, 1981; Swinney & Prather, 1989; Holcomb, Coffey & Neville, 1992). Earlier studies have used such direct methods of a child's interpretation of a sentence as the picture selection task, the act-out task, or the truth-value judgment task (McDaniel, McKee & Cairns, 1996) that presumably provide an off-line measure of comprehension processes (McKee, 1996). Children's on-line comprehension is only now becoming the focus of empirical and systematic investigation. Booth *et al.* (2000) and Traxler (2002) employed self-paced reading and listening tasks to examine how eight-to-twelve-year-old children process unambiguous, but complex sentences with relative clauses. They found that children's comprehension of such sentences is much less accurate and more laborious than that of adults. Hahne & Friederici (2001) provided neurolinguistic evidence from Event-Related Potentials that seven-to-eight-year-old children process sentences in a fashion similar to adults, but take longer to comprehend syntactically anomalous ones. Trueswell *et al.* (1999) were the first to examine whether much younger, pre-literate four-to-seven-year-olds employ syntactic and referential factors on-line in order to resolve syntactic ambiguity. The head-mounted eye-tracking system allows for monitoring children's eye movements while they listen to the spoken instructions, a paradigm that does not rely on children's yet to be developed reading skills. The studies that followed have established that the head-mounted eye-tracking technology can successfully tap not only into children's real-time processing of referential ambiguity, but also grammatical gender (Arnold, Brown-Schmidt, Trueswell & Fagnano, in press), contrastive focus (Nadig & Sedivy, 2000), and verb lexical bias (Snedeker, Thorpe & Trueswell, 2001).

The two eye-tracking studies that follow contribute to the ongoing research on adults' referential ambiguity resolution (Experiment 1) and provide novel evidence on how children look for referents for ambiguous pronouns in a real-time fashion (Experiment 2). We took Runner *et al.*'s (in press) findings as reasonable evidence for feasibility of the eye-tracking methodology to assess listeners' on-line referential ambiguity resolution strategies. Following Piñango *et al.* (2001), we considered pronouns and reflexives to induce comparable processing load and hypothesized that longer reaction times and eye gaze latencies for pronouns compared to reflexives should be attributed to referential ambiguity. Experiment 1 consisted of an off-line referent preference questionnaire and an eye-tracking study. The off-line questionnaire establishes the baseline for choice of the referent and empirically tests the linguistic hypothesis about referential ambiguity of SDPs. The on-line eye-tracking study provides evidence from the reaction times data for the discourse processing complexity hypothesis (Piñango *et al.*, 2001), and investigates the moment-by-moment nature of the referential ambiguity resolution. Experiment 2 is an eye-tracking study



A. *The boy has placed the box behind himself/him.*

B. *The woman has placed the horse near herself/her.*

Fig. 1. Examples of stimulus pictures (Trials 27 and 16).

with children that compares on-line measurements (eye movements) and off-line behaviour data in the picture selection task.

EXPERIMENT 1

PROCESSING OF REFERENTIALLY AMBIGUOUS PRONOUNS BY ADULTS

Before conducting the eye-tracking study, we established the baseline for choosing a referent for SDPs in an off-line paper-and-pencil questionnaire. In this experiment, we examined referent preferences for both reflexives and pronouns for adult English speakers and investigated empirically the claim that SDPs are referentially ambiguous.

REFERENT PREFERENCE QUESTIONNAIRE

Participants

Fifty-eight participants, all native English speakers from the Rutgers University undergraduate population, took part in the experiment, 29 in each of its two versions. They were naïve with respect to the purpose of the experiment. Participants received course credit for participation. Completion of the questionnaire typically took 20 minutes.

Design and materials

Participants were presented with 24 pairs of pictures (Figure 1).[1] Each picture in a pair contained two characters and an object. The pictures were identical except for the location of the object and some other visual details.

[1] The pictures for all the experiments reported in this article were modified from the original pictures used in a series of Norwegian experiments in Hestivk & Philip (1999/2000). The authors would like to thank Ned Norland for his assistance with modification of the original pictures.

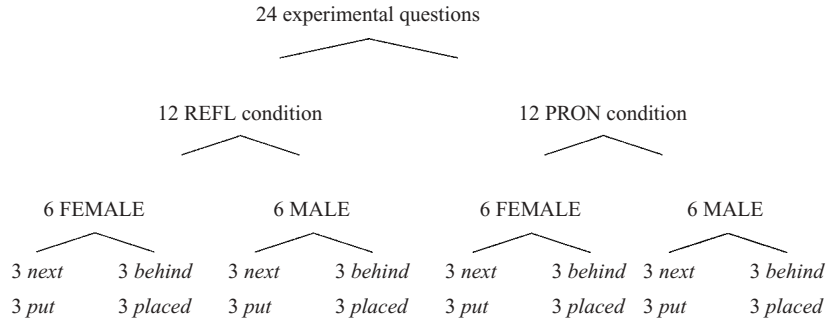


Fig. 2. The factors manipulated in the experimental items.

The pairs of pictures always contrasted either two male characters (a boy, a man, or a grandfather) (Figure 1A) or two female characters (a girl, a woman, or a grandmother), as in Figure 1(B). The orientation of the character with respect to the object was either facing the object, with the preposition *near*, or looking away from the object, with the preposition *behind*. Twelve different objects were used: bucket, box, hairbrush, horse, teddy bear, pig, glass, hat, suitcase, ball, chair, and watering can. Each object appeared twice in the course of the experiment. The linear order of the pictures in the pair (left vs. right), the location of the object, the linear order of the two characters, and their orientation were counterbalanced.

Three factors were manipulated in the experimental materials: the gender of the two characters (female vs. male), the verb (*put* vs. *placed*), and the preposition (*behind* vs. *near*). The experimental questions were constructed as 24 pairs; questions within the pair differed only in whether they contained a PP with the reflexive (reflexive condition) or with the pronoun (pronoun condition). The questionnaire was assembled according to a design schematically represented in Figure 2. The inclusion of these factors, gender, verb, and preposition, was necessary to check whether idiosyncratic lexical and pragmatic factors contribute to the referent preference pattern.

So that identical materials would not be repeated to any participants, the questionnaire was constructed in two versions, with the reflexive and pronoun conditions distributed over versions in a counterbalanced design. There were no fillers in the questionnaire. Each of the 24 picture-pair-plus-text items consisted of two preamble sentences, an experimental question, and three possible answers, as illustrated in (6).

(6) PREAMBLE: *In these pictures, you see a boy, a man, and a box. The boy has placed the box on the ground.*
REFLEXIVE QUESTION: *Which picture shows that the boy has placed the box behind himself?*

PRONOUN QUESTION: *Which picture shows that the boy has placed the box behind him?*

ANSWERS: (*a*) *the left picture* (*b*) *the right picture* (*c*) *both pictures*

In contrast to the verb, gender, and preposition manipulations, the accessibility of the referent for the pronoun was held constant across the conditions. We followed Arnold *et al.* (2000) in judging the sentence-internal referent, *the boy* in (6), as highly accessible in spoken discourse due to its salience (mentioned three times) and recency. In contrast, the sentence-external referent, *the man*, was not as accessible as *the boy* because it was mentioned only once and was separated from the pronoun by two sentences. Visually, however, both referents were equally salient.

Procedure

Each version of the questionnaire was presented to a different group of participants in a paper-and-pencil format. Participants were instructed to examine the pictures, read the preamble sentences and the question carefully, and then indicate their answer by circling one of the three possible answers. The responses in the reflexive condition were screened for errors, and two participants with 50% or more errors (six out of 12 items) were rejected.

Data treatment

Participants' answers were recorded as mean percentages of sentence-internal and sentence-external referent preferences in two matrices, one for each of the two conditions, reflexive vs. pronoun. The data were analysed parametrically in analyses of variance, with the percentage of sentence-internal preference as the dependent variable. Additional analyses were conducted to look for effects of lexical (preposition and verb) and pragmatic (gender) factors.

The terms of the ANOVA bear directly on the question around which the experiment was designed. The main effect of the reflexive vs. pronoun factor provides information about whether participants gave reliably different referent preference for the questions with the pronoun in contrast to the questions with the reflexive. The statistical significance of this term indicates an overall preference for the sentence-internal referent interpretation in the reflexive condition, presumably reflecting the ambiguous status of the pronoun in the pronoun condition. The non-significance in the (2) preposition × (2) verb × (2) gender × (2) pronominal type interaction term of the ANOVA indicates a reasonable consistency across questions in selecting the sentence-internal referent interpretation. This means that idiosyncratic

TABLE 1. *Adult referent preference questionnaire* (N = 56)*: overall sentence-internal referent preference* (%)

	Percentage referent choice for the pronominal		
	Sentence-internal	Sentence-external	Both
REFL	95**[b]	5[a]	0[a]
PRON	79**[b]	17	4

[a] These would be incorrect reponses.
[b] $p < 0{\cdot}001$.

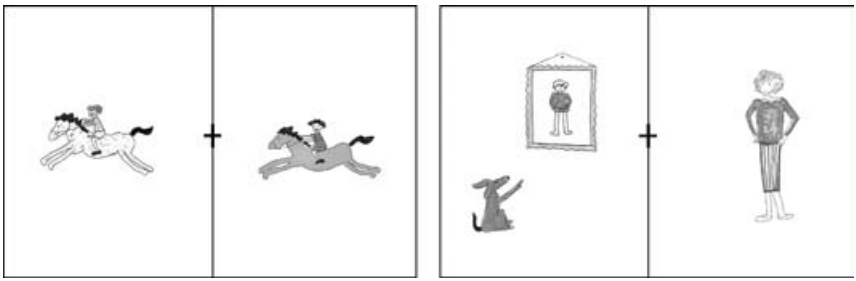
lexical and pragmatic factors did not contribute to the referent preference pattern.

Results

Table 1 reports the percentage of sentence-internal and sentence-external referents for the two pronominal types.

The reflexive questions were overall interpreted correctly: the participants chose the sentence-internal referent in 95% of the cases. That is, in Item 27 with the reflexive (Figure 1), they chose Answer (a) 'the left picture' indicating that *himself* takes the sentence-internal referent *the boy* as its antecedent. The responses in the pronoun questions reflect the ambiguous referential status of the pronoun; overall, in 21% of cases, the participants chose the subject-external referent for the ambiguous pronoun. That is, in Item 27 with the pronoun, participants chose Answer (b), 'the right picture', or sometimes (c), 'both pictures', indicating that *him* can take the sentence-external referent *the man* as its antecedent. The difference in preference for sentence-internal referent for reflexive vs. pronoun was highly significant, ($F1(1, 57) = 24{\cdot}11$, $p < 0{\cdot}001$, $F2(1, 23) = 45{\cdot}16$, $p < 0{\cdot}001$). There were no main effects or interactions for any of the three lexical factors: verb, gender, or preposition.

In sum, these results showed that adult English speakers preferred the sentence-internal referent for the pronoun. This is expected if salience and recency-of-mention play a crucial role in the referent accessibility. What is remarkable is that the participants nevertheless chose the sentence-external interpretation for the pronoun in 21% of the cases, despite the fact that the sentence-external referent was presumably more difficult to access. In Arnold *et al.*'s (2000) study, an addition of just one sentence between the pronoun and the referent (Example 5) made the correct sentence-external interpretation highly infelicitous and resulted in listeners choosing an incorrect but more accessible referent. Thus, adult English speakers in the referent preference questionnaire indeed found the sentences with SDPs ambiguous, supporting the claim made in the linguistic literature.



A. *Which picture shows that the boy is riding a brown horse?*

B. *Which picture shows that the dog was pointing to him?*

Fig. 3. Examples of filler pictures (Trials 12 and 2).

ADULT EYE-TRACKING EXPERIMENT

The questionnaire study provided evidence that adults noticed the referential ambiguity in the SDP context but preferred the sentence-internal referent for the ambiguous pronoun. How do participants make this choice? The off-line nature of the questionnaire did not provide any information about how this referential ambiguity was considered and resolved. Experiment 2 investigated the moment-by-moment ambiguity resolution in an eye-tracking study.

Participants

Eighteen participants, all native English speakers from the Rutgers University undergraduate population, took part in the experiment, nine in each of the two versions of the experiment. They were naïve with respect to the purpose of the experiment. Participants received part of course credit for their participation. Completion of the experiment typically took 40 minutes.

Design and materials

The materials included the 24 experimental items from the questionnaire and, in addition, 30 fillers. The fillers also always contained a pair of pictures, sometimes very similar to each other (Figure 3A), so as to resemble the experimental trials, and sometimes quite different (Figure 3B).

The instructions in this experiment consisted of four-to-five sentences: two preamble sentences (an introduction and the fixation instruction), the experimental question, and one or two follow-up instructions. Example (7) illustrates the experimental item 27 (Figure 3A), and Example (8) a filler (Figure 3B).

(7) PREAMBLE: *In these pictures, you see a boy, a man, and a box.*
FIXATION: *Now look at the cross.*

Fig. 4. Remote tabletop ISCAN eye-tracking system.

(8)
QUESTION: *Which picture shows that the boy has placed the box behind himself/him?*
FOLLOW-UP: *Now look at the boy with the cap on.*
PREAMBLE: *In these pictures, you see a dog, a painting, and a woman.*
FIXATION: *Now look at the cross.*
QUESTION: *Which picture shows that the dog was pointing to him?*
FOLLOW-UP: *Now look at the dog's tail.*

Note that the accessibility of the two potentials referents was different, although the sentence-external referent, *the man*, was potentially more accessible in (7) than in (6). Only one sentence separated the pronoun and the referent, and it did not contribute to the proposition. The sentence-internal referent, *the boy*, remained highly accessible. The same experimental manipulations as in the questionnaire (i.e. gender, verb type, and preposition type) were used. Two versions of the experiment were prepared in a counterbalanced design, with 54 items in each.

Procedure

The current experiment used a remote tabletop ISCAN eye-tracking system designed to provide maximum freedom for subjects. This system consisted of an eye-tracking computer, a remote camera, and a stimuli presentation computer (Figure 4).

The eye-tracking computer controlled the remote camera positioned on a tilting pan that recorded a close-up image of the left eye. The pan was located in front of the participant on a desk at a distance of approximately two feet. The stimulus pictures were presented on a computer screen

positioned immediately behind the remote camera pan. The computer screen image was converted from PC format to video format with the help of a PCI card. The converted signal was fed to the eye-tracking system, which projected it onto a separate screen monitor with a superimposed eye position marker (a cross-hair) not visible to participants.

Because the eye camera was remote, movement of the participant's head was restricted to the range of 12 inches horizontally and four inches vertically. Head movements resulted in the automatic repositioning of the tilting pan, and the camera followed the eye within this range. Eye position continued to be plotted on the scene monitor throughout any movement of the head. The stimulus image and the superimposed eye position, along with all auditory stimuli, were recorded on tape using a frame-accurate digital video recorder (a SONY DSR-30).[2] Prior to the experiment, participants underwent a calibration procedure. During this time, an experimenter obtained a participant's eye image by manually adjusting several components of the eye-tracking system to provide the eye-tracking computer with alignment information.

Participants were tested individually, with an experimental procedure that involved two tasks. They looked at the pictures and made a decision about which of the two pictures represented the correct answer to the experimental question. A button-box with left and right buttons was used for the forced-choice task. The visual stimuli, that is, the pairs of pictures, were presented on a 19-inch Dell desktop monitor, under the control of DMDX software (Forster & Forster, 2003). Each display consisted of two pictures of equal size separated by a line with a cross-hair in the middle (see Figures 1 and 3), with the pictures filling up the entire screen. Participants controlled the pace of the picture presentation by pressing the item request button after they had made their picture choice.

A female experimenter gave the verbal instructions live, pronouncing them with a normal speed. The same female voice was used for all participants. Once the stimulus pictures appeared on the screen, the instructions began. Instructions always began with the preamble sentence (i.e. *In these pictures, you see a boy, a man, and a box*) and a request to look at the central fixation point (*Now look at the cross*), followed by the experimental question and one or two additional instructions to look at various parts of the pictures. The experimenter who gave the verbal instructions paused for a short time after the question to make sure that the participant had time to choose one of the two pictures by pressing the appropriate button. She then said 'Next' and proceeded to the remaining instructions. Prior to analysis, we selected a random sample of utterances for acoustic analysis. A trained

[2] We refer the reader who is interested in a more detailed description of the ISCAN eye-tracking system to Trueswell *et al.* (1999: 98–100).

phonetician judged the reflexive and pronoun utterance samples to be comparable both prosodically and acoustically. That is, the placement of intonation phrase boundaries and shape and placement of pitch accents were comparable across conditions, and the overall pitch range, amplitude and speaking rate were also similar. Participants were instructed to listen carefully to the instructions and to very briefly fixate on the cross during each trial. They were then to press the appropriate button as quickly as possible to indicate their choice of the picture, and continue examining the pictures until they heard the instruction 'Next'.

Data treatment

The response data in the reflexive condition were screened for errors, and no one was rejected according to this criterion. In addition, the eye-tracking data were checked for track loss. Two participants whose eye-tracking data were judged to be poor due to 25% track loss were rejected, resulting in the total of 16 participants for the final data analysis.

Three types of data were analysed in this experiment: responses in the forced-choice task, reaction times for the forced-choice task, and eye movements. First, just as in the questionnaire, the response data for the mean percentages of the sentence-internal referent preference were assembled into two matrices, one for each of the conditions, reflexive vs. pronoun and were analysed parametrically in analyses of variance. Second, the reaction time data were analysed for information about whether participants took longer to choose a picture in the pronoun condition compared to the reflexive condition. The reaction times were calculated from the onset of the pronominal in the sentence, that is, from the onset of *him* or *himself*, until the participant pressed the button.[3] Statistical significance in this analysis indicates that longer reaction times in the pronoun condition reflect the ambiguous status of the pronoun, which requires additional processing. Finally, the eye movement data showed the probability of fixating on each picture reflecting either sentence-internal or -external referent preference during the course of interpreting spoken sentences, and provided information about how the ambiguity is resolved in real-time.

Since the participants' task was to perform a speeded force-choice task, i.e. to press the right or left button to indicate their choice of the two pictures, we would expect the majority of button presses to occur after the spoken instruction was completed. However, as is typically the case with the reaction time data, in a certain number of trials, the subjects pushed the

[3] The raw reaction times that were faster than 100 ms and slower than 3500 ms were adjusted to the appropriate subject- and item-based means; these data points accounted for 5·2% of the collected RT data.

TABLE 2. *Adult eye-tracking experiment 2 (N = 16): overall sentence-internal referent preference (%)*

	Percentage referent choice for the pronominal	
	Sentence-internal	Sentence-external
REFL	97**[b]	3[a]
PRON	80**[b]	20

[a] Incorrect responses.
[b] $p < 0{\cdot}001$.

button ahead of time, without waiting until the end of the question. Once the button was pushed, the trial was considered completed and recording of eye movements stopped. We had the 33% percentages of trials completed before the pronominal occurred in the question and 63% after. Thus, the discussed below are based on approximately two-thirds of the collected data. Since the missing data present a problem for the balanced analyses of variance for many statistical programs, we performed our analysis using the SAS statistical software that makes it possible to test a statistical hypothesis based on imbalanced data.

Results

Responses. The preferences of sentence-internal and sentence-external referents for the two pronominal types in the response data obtained in the eye-tracking study closely resemble those from the questionnaire (Table 2).[4]

Participants chose the sentence-internal interpretation statistically significant higher percent of the time for reflexive than for pronoun ($F1(1, 15) = 4{\cdot}63$, $p < 0{\cdot}05$, $F2(1, 23) = 46{\cdot}79$, $p < 0{\cdot}001$). There were no significant main effects or interactions for gender, verb, or preposition. Thus, the on-line results of the eye-tracking study replicated those of the off-line questionnaire: the pronoun sentences are ambiguous for adults, although they do prefer sentence-internal referents.

Reaction times. Table 3 below summarizes the reaction time data, for each of the two conditions.

Since the previous research (Runner *et al.*, in press; Piñango *et al.*, 2001) has demonstrated that pronouns and reflexives induce comparable processing load, we suggest that the participants' statistically significant slower reaction times for the pronoun sentences reflect ambiguity resolution absent

[4] Since the task was forced-choice, only two types of responses were possible: the sentence-internal interpretation and the sentence-external interpretation. In the questionnaire, the additional option was 'both'.

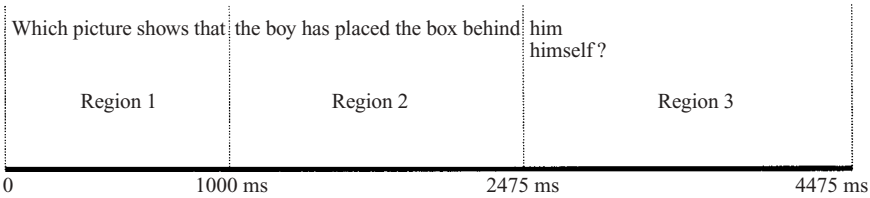


Fig. 5. The average speech onsets of the three regions in the experimental items.

TABLE 3. *Adult eye-tracking experiment* 2 (N = 16)*: reaction times* (*ms*)

	Reaction times
REFL	1264
PRON	1430*[a]

[a] $p < 0{\cdot}05$.

in the reflexive sentences ($F1(1, 14) = 4{\cdot}25$, $p = 0{\cdot}058$, $F2(1, 22) = 4{\cdot}64$, $p < 0{\cdot}05$).

Eye movements. Prior to the eye movement analysis we calculated the onsets of the three regions in the sentences (Figure 5).

Region 1 started at the onset of *which* in *Which picture shows that* ... and was always set to zero. The duration of Region 1 was remarkably consistent, 1000 ms on average (± 33 ms). Region 2 started at the onset of the embedded clause and continued until the pronominal element, i.e. *the boy has placed the box behind* ... . The duration of this region varied depending on the length of the four types of lexical items: the agent noun, the verb, the object noun, and the preposition.[5] Finally, Region 3 started at the onset of the pronominal element (at 2475 ms on average) and continued until the onset of the second sentence in the instructions. On average, the reflexives, *himself* and *herself*, were approximately three times as long as the pronouns, *him* and *her* (330 ms vs. 100 ms).

Since the sentences in the two conditions were identical until the pronominal element, we coded the video recording of participants' eye movements in Region 3 only, beginning at the onset of *him* or *himself*. For each 33-ms video frame, we performed a coarse-grain analysis of eye movements by identifying which of the two pictures the participant was fixating. Thus for Item 27, *The boy has placed the box behind him* (Figure 1),

[5] There were six agent nouns in total: *boy, girl, man, woman, grandfather, and grandmother*. On average, the long agent nouns, *grandfather* and *grandmother* were twice as long as the short ones, *boy* and *girl* (650 ms vs. 330 ms). Similar duration variability was found for the object nouns (*box, cup* vs. *hairbrush, watering can*).

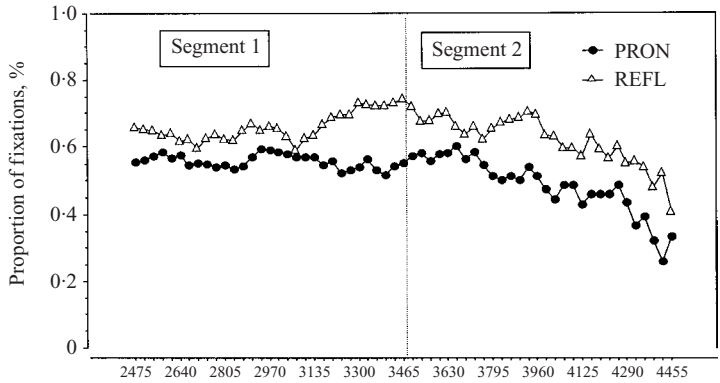


Fig. 6. Adult eye-tracking experiment 2: probability of fixating the sentence-internal referent picture in two conditions (%), for the entire Region 3.

the significantly greater number of looks to the left picture indicates the sentence-internal interpretation and to the right one the sentence-external interpretation.

Figure 6 presents the graphical results of the coarse-grain analysis. It shows the proportion of fixations in percent on a sentence-internal picture in the pronominal and reflexive conditions for each 33-ms frame in Region 3, starting at the onset of *him/himself* (75 frames = 2475 ms).

There were significantly fewer looks to the sentence-internal picture in the pronoun condition than in the reflexive condition reflecting a stronger competition between the sentence-internal and -external referents. This indicates that participants were aware of the referential ambiguity in the pronoun condition and were resolving it on-line.

To test the reliability of looks to the sentence-internal vs. sentence-external picture, we performed analyses of variance in three different ways: on the entire Region 3 and also by splitting this region into two equal segments. The first segment started at the onset of the pronominal and lasted for 1000 ms (2475–3474 ms). The second segment began 1000 ms after the onset of the pronominal (3475–4475 ms) and also lasted for 1000 ms (indicated in Figure 6 by the dotted line). Conducting analyses on two separate segments allowed us to identify the real-time course of the on-line resolution of referential ambiguity. For the entire Region 3, ANOVAs revealed that the adults consistently looked significantly more at the sentence-internal picture in the reflexive condition than in the pronoun condition, 65% vs. 55% ($F(1, 15) = 5{\cdot}71$, $p < 0{\cdot}05$). The same pattern was found for Segment 1, main effect of condition, 63% vs. 55% ($F(1, 15) = 4{\cdot}83$, $p < 0{\cdot}05$), and Segment 2, 69% vs. 55% ($F(1, 15) = 4{\cdot}85$, $p < 0{\cdot}05$). There were no other significant main effects or interactions.

Thus, the coarse-grain analysis of eye movements showed that the adult participants were able to detect the referential ambiguity in the sentences with pronouns as soon as they heard the pronoun, and rapidly took into consideration the alternative possible interpretations. The resolution of the referential ambiguity took place soon after the lexical information about the pronominal became available, around 250–350 ms after the onset of the pronominal *him/himself*. This time course is comparable to rapid and incremental referential interpretation of lexical NPs (Tanenhaus *et al.*, 1995; Trueswell *et al.*, 1999): when a single unique referent is present in the scene, adults launch eye movements to the intended referent within 250–350 ms of perceiving the noun.

To summarize so far: the results of the eye-tracking study replicated the results of the questionnaire study. For adult English speakers, all three types of analysed data – responses, reaction times, and eye movements – indicate that SDPs are referentially ambiguous. Moreover, neither the difference in accessibility of the two possible referents nor the necessity to identify the correct referent under pressure decreases the roughly 20% choice of the sentence-external referent. Adults' processing resources are flexible enough for them to be capable of making inferences on-line with respect to less accessible referent, showing a competition between the two referents shortly after encountering the pronoun in the sentence. However, the adults did take longer to access the sentence-external referent, resulting in slower reaction times in choosing a picture for sentences with the pronoun. This finding supports the discourse processing complexity hypothesis by Piñango *et al.* (2001).

Is it the case that due to children's limitations on processing and inferential resources, they will fail to resolve this referential ambiguity? We address this question in a children's eye-tracking experiment.

EXPERIMENT 2

CHILDREN'S EYE-TRACKING EXPERIMENT

Participants

The participants were sixteen children aged 4;9 to 7;10 attending either kindergarten, the first or the second grade (mean age 6;6, eight 4;9 to 6;5 and eight 6;9 to 7;10). As with the adults in Experiment 2, eight children were assigned to one version of the experiment and eight to the other. Children received a toy for their participation. The study was conducted on the schools premises, and typically took 20 minutes.

Design and materials

The target materials and design were basically the same as in Experiment 1, with several methodological modifications. First, the spoken instruction

sentences were pre-recorded by the same speaker. Second, to make sure that children could attend to the task, the children's version of the experiment was shorter. They saw 16 experimental items and 16 fillers out of the 54 items (24 and 30, respectively) in the adult experiment. The instructions consisted of four sentences: the preamble, the experimental question, and one follow-up instruction. The child's task was also a forced-choice picture selection task. However, the children did not have to press buttons; instead, they had to choose a picture by pointing to it with their finger. Thus, although the picture selection task provides a relatively direct measure of a child's sentence interpretation it only taps later, off-line stage of sentence comprehension and there is debate about the consistence of the effects found with such tasks (Gerken & Shady, 1996).

The instructions were adapted to accommodate pointing instead of looking, as illustrated in (9; *cf.* (7)) for experimental item 27 (see Figure 1A).

(9) PREAMBLE: *In these pictures, you see a boy, a man, and a box.*
FIXATION: *Now look at the cross.*
REQUEST: *Now point to the picture where the boy has placed the box behind himself/him.*
FOLLOW-UP: *Now point to the boy with the cap on.*

The same experimental manipulations from Experiments 1 and 2 (gender, verb type, and preposition type) were used in Experiment 2. Two versions of the experiment were prepared in a counterbalanced design, with 32 items in each.

Procedure

While the adult Experiment 1 used a remote tabletop ISCAN eye-tracking system, an ISCAN head-mounted eye-tracking visor was used for the children. The former is designed to be used with electronically presented materials while the latter with real objects. Thus, the adults in Experiment 1 saw the pictures presented on the computer screen and the children in Experiment 2 saw the same pictures on paper. With the electronic presentation, the picture fills in the entire screen: no background is recorded and the picture is kept stable making eye movement coding easier. With the paper presentation, the eye-tracking system records the picture and its surroundings; how much depends on how close the participant is to the picture. It also records changes in the picture's orientation resulting from the participant's tilting the head during the experiment. While it would be easier for subsequent coding to present the materials to children electronically as well, we had to switch to the head-mounted version of the eye-tracker. There were three practical reasons for doing so. First, with the remote eye-tracking system, movements of the participant's head were

restricted to the range of 12 inches horizontally and four inches vertically. When we piloted the remote system with two children, we found it was too difficult for children to restrict their body and head movements within the required range. Second, the child participants for the experiment were recruited and run at local schools. Since the remote ISCAN eye-tracking system is stationary, we used the portable head-mounted eye-tracking system. Finally, children in the pilot experiment often poked the delicate LCD laptop screen with their finger despite regular gentle reminders just to point, not to touch. In our view, these minimal practical differences between Experiments 1 and 2 do not present a problem for comparison of the results. The two systems have the same technical characteristics, record eye movements in an identical fashion and can be used interchangeably.

The ISCAN head-mounted eye-tracking visor used with children consisted of a monocle and two miniature cameras. One recorded the scene from the participant's perspective and the other recorded an image of the left eye. The cameras were attached to the visor that was worn like a cap, and the children's natural body and head movements during the experiment did not present a problem for the accuracy of the eye movement data. Children were run individually, in a quiet room at the school they attended. One research assistant presented the experimental pictures in hard copy to the child and interacted with her during the study. A second research assistant played the sound files on the computer at a rate appropriate for each child. Finally, the third experimenter experienced in operating the eye-tracker evaluated the picture image and the eye position and adjusted the angle of the eye-tracking monocle as needed during the experiment. Prior to conducting the experiment, each child participant was familiarized with the experimenters, the equipment and the task requirements. After obtaining the child's oral consent, the visor was positioned on her head, and an eye-tracking calibration procedure was performed.[6] The child was asked to briefly hold her gaze on a sequence of five spatial positions on an $8' \times 11'$ page; each position contained a coloured picture of a familiar animal. Our success rate for achieving a fixed head position was high, resulting in accurate calibration of the eye-tracker, usually on the second attempt.

The child then was seated at a child-size table, and a binder with the pictures was positioned vertically on the table at the child's arm reach. The first experimenter was seated opposite to the child so that she could flip the pages, monitor the child's behaviour during the experiment and provide encouraging feedback. The second experimenter was at a different table behind the child from which she could see her pointing actions and play spoken instructions from the laptop computer. The speakers were located

[6] A written consent form was signed by each child's parent.

on the child's table, on either side of the binder. The third experimenter and the eye-tracking equipment cart were also stationed behind the child, as far away as the cable connecting the visor and the eye-tracker would allow.

The pictures were printed in colour on paper and assembled in a binder. Each trial such as (9) started when the first experimenter flipped the page in the binder and the preamble sentence (*In these pictures, you see a boy, a man, and a box*) was played. The child then heard the fixation instruction (*Now look at the cross*). The sound presentation of the experimental instruction that required pointing was delayed until the third experimenter confirmed that accurate cross fixation had been obtained. Upon hearing the experimental (*Now point to the picture where the boy has placed the box behind him*) and the subsequent follow-up instructions (*Now point to the boy with the cap on*), the child proceeded to point to appropriate parts of the picture with her finger. After each trial children were given encouraging feedback.

Data treatment

Only two types of data were analysed in this experiment: pointing responses in the picture selection task and eye movements. As in Experiment 1, the response data were assembled into two matrices, one for each of the conditions, reflexive vs. pronoun, for parametric analyses of variance. The reaction times that resulted from pointing were not collected due to the highly subjective nature of such actions that depended on each child's individual motor skills development. For the eye movement data, we calculated the probability of fixating on each picture during the course of interpreting spoken sentences, which provided information of how the ambiguity is resolved in real-time.

Results

Responses. The preferences of sentence-internal and sentence-external referents for the two pronominal types in the off-line response data obtained for children in Experiment 2 are quite different from the referent preferences for adults reported in Experiment 1 (Table 4).

Children's referent preferences were identical in both reflexive and pronoun conditions ($Fs < 1$): they overwhelmingly chose the sentence-internal referent (*the boy*). The 7% sentence-external preference is nearly the same as the percent of errors in the reflexive condition, suggesting that it cannot be interpreted as a meaningful choice. There were no significant main effects of gender, verb, or preposition. Thus, judging from the response data only, one might infer that while the pronoun sentences are ambiguous

TABLE 4. *Children eye-tracking experiment 3 (N = 16): overall sentence-internal referent preference (%)*

	Percentage referent choice for the pronominal	
	Sentence-internal	Sentence-external
REFL	94**[b]	6[a]
PRON	93**[b]	7

[a] Incorrect responses.
[b] $p < 0{\cdot}001$.

for adults, they are totally unambiguous for children, with the sentence-internal referent as the only possible antecedent for the SDP. This conclusion, however, is not supported by the eye movement data.

Eye movements. As with the adult eye-tracking data from Experiment 1, the experimental sentences were divided into three regions (see Figure 5 above) with Region 3 being the region of interest. In order to make a direct comparison with the adult eye movement data, the stop point at 4475 ms was used. Note that it took the children longer to do the pointing: in 66% of the trials the children had pointed to one of the pictures by the stop point, while in 96% of the trials the adults had pushed the button. We did not analyse the right tail of the eye movement distribution because of great variability in the remaining data. For the coarse-grain analysis, we analysed the proportion of fixations (in percent) on the picture consistent with the sentence-internal interpretation compared to the picture with the sentence-external interpretation.

Figure 7 presents the percentage of fixations on either the sentence-internal or sentence-external picture for the entire Region 3.

Taken as a whole, there was no significant difference between the looks to the sentence-internal picture in the sentences with reflexives than with pronouns, 61% vs. 56%. However, we found a quantitatively different pattern in children's moment-by-moment processing compared to that of the adults if we inspect each of two segments separately. In Segment 1 (2475–3475 ms), there still was no difference in the children's looks to the sentence-external picture in the reflexive condition compared to the pronoun condition (57% vs. 55%, $F < 1$). But crucially, Segment 2 (3475–4475 ms) revealed a significant adult-like difference between the reflexive and pronoun conditions: the children indeed looked substantially less at the sentence-internal picture in the latter, 62% vs. 43% ($F(1, 15) = 11{\cdot}33$, $p < 0{\cdot}01$). There were no other significant main effects or interactions. We interpret these results as an indication that the children take much longer than adults to notice the ambiguity, thus, no difference in looks between the two pictures for the first 1000 ms after the onset of the

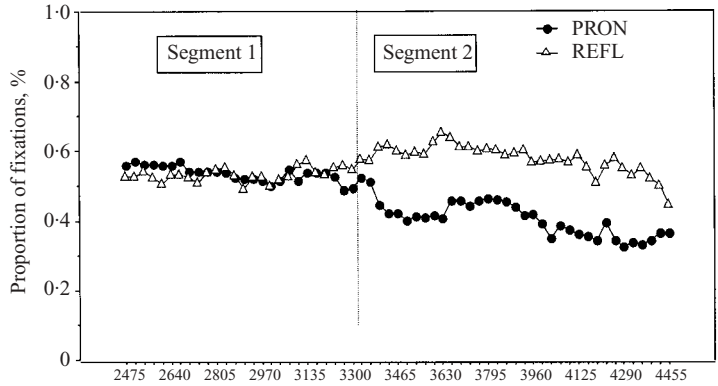


Fig. 7. Children's eye-tracking experiment 3: probability of fixating the sentence-internal referent picture in two conditions (%), for the entire Region 3.

pronominal. However, closer towards the end of the trial, during the second 1000 ms, the children start to notice the referential ambiguity in the pronoun condition. Crucially, although their eye movements show that they are becoming aware of this ambiguity, it did not affect their final choice of the antecedent. The pointing data from Table 4 above testify to the fact that the children did not revise their initial commitment to choose the sentence-internal referent despite the fact they were considering the sentence-internal referent significantly less in the sentences with the ambiguous pronouns.

In sum, comparison of eye movements of adults and children revealed differences in the timing of pronoun ambiguity resolution. Adults quickly notice the referential ambiguity and start resolving it on-line, within 1000 ms after the onset of the pronoun. Their response data show 80% choice of the sentence-internal referent and 20% choice of the sentence-external referent for the ambiguous pronoun condition. Children, on the other hand, are much slower in noticing the ambiguity; during the initial 1000 ms after the onset of the pronoun, they consider the sentence-internal interpretation. The situation changes later in the trial as they become aware of two potential choices for the referent. But their response data do not give any hint that this is the case. The overwhelming 93% choice of the sentence-internal referent for the pronoun showed children's reluctance to revise their initial commitment to the single interpretation revealing their deterministic approach in resolving referential ambiguity.

GENERAL DISCUSSION

The two adult experiments produced three main results. First, we have found empirical evidence that adults do find the pronouns in the short-distance contexts ambiguous, supporting the claims made by the current

processing literature. While the grammar permits sentence-external reference and, as an exception, sentence-internal one, and the visual context is consistent with both, adults readily access the latter. The preference for the sentence-internal reference has a performance-based explanation since it has no plausible basis in grammar *per se*. According to the performance-based account, the sentence-internal referent is favored by the 'minimum commitment' strategy of the language processor: accessibility, order-of-mention, recency, and the thematic role of the sentence-internal referent make it the default option. In contrast, the performance system makes the sentence-external referent challenging, as was demonstrated by the increased reaction times when the adult participants were processing referentially ambiguous pronouns. This finding provides evidence in support of the discourse processing complexity hypothesis argued for from the theoretical perspective by Reinhart (1999) and experimentally by Piñango *et al*. (2001). Reinhart proposes an explanation in terms of semantic processing: the sentence-internal reference is established via variable-binding and is less costly since it enables immediate closure of open properties. The sentence-external reference is a case of coreference ('accidental' in terms of Thornton & Wexler, 1999), it is only allowed in a restricted range of discourse contexts and requires that the property is stored open until an antecedent for the pronoun is found. Similarly, Piñango *et al*. showed that the interpretative processes that demand access to discourse information during real-time comprehension are costly. Processing of reflexives, where the interpretation is obtained within the clause via syntactic mechanisms alone, is less costly than that of referentially ambiguous pronouns. In the latter case, the processor has to access information beyond that provided in the syntax, i.e. discourse.

Second, our experiments demonstrate that the eye-tracking method is a feasible probe for testing binding theories that is especially suited to accessing speakers' intuitions under complex discourse conditions (see also Runner *et al.*, in press). Recording of eye movements allows for better understanding of what people are doing while resolving referential ambiguities on-line. Previous eye-tracking studies (Tanenhaus *et al.*, 1995; Trueswell *et al.*, 1999) found evidence for rapid incremental referential interpretation of lexical NPs. When a single unique referent is present in the scene, adults launch eye movements to the intended referent within 250–350 ms of perceiving the noun. In our study, we found that adults engage in anaphoric resolution (that is, look to link the pronoun with a potential antecedent) within 500 ms from the onset of the pronoun, and continue resolving the referential ambiguity throughout the trial.

Finally, our results from the adult experiments support a constraint-based model of language processing (Tanenhaus & Trueswell, 1995, among others), in which multiple sources of information are used on-line to guide

referential processing. Visual context appeared to be a compelling source of information. It kept both the sentence-external and -internal referents activated throughout the trial and made the more difficult sentence-external referent a viable competitor in the referential ambiguity resolution. Note that this finding is in contrast to Arnold *et al.*'s (2000) results: we found that the choice of the sentence-external referent was not affected by its low accessibility, as it was chosen a robust 21% of the time in both experiments.

How did children's processing of the referentially ambiguous pronouns differ from that of adults? As was expected, the behavioural data from the picture selection task indicate that children overwhelmingly prefer the sentence-internal referent. This preference can be readily explained by the fact that the adults preferred the sentence-internal referent, and the children used the same default as adults. In contrast to the adults, however, the children in our study did not overtly select the sentence-external referent in the sentences with referentially ambiguous pronouns supporting the pattern of difficulty well-attested in the acquisition literature. Avrutin (1994) argues that children face difficulties in all areas of anaphora resolution in discourse. Thornton & Wexler (1999) suggest that the ability to access the sentence-external interpretation depends on children's knowledge of discourse 'guises' which have yet to be acquired. Other child experiments testing the interpretation of empty categories in a variety of contexts also demonstrate a difficulty with external reference (Goodluck, Terzi & Chocano Díaz, 2001). Since the difficulty with the sentence-external referent and preference for the sentence-internal one is often viewed as a consequence of the structure of the sentence processing mechanism in general, this effect is exacerbated in children due to their quantitatively smaller processing capacity.

The notion of the child's underdeveloped processing mechanism has been proposed to account for many cases of nonadult behaviour in establishing pronominal reference. Grodzinsky & Reinhart (1993) proposed that the much greater failure of children than adults in Principle B contexts arises from the child's processor being unable to consider two representations at once, the bound variable and the accidental coreference representation. Computing the accidental coreference that requires a child to construct two representations, keeping both in working memory, and then selecting the appropriate one, exceeds the child's processing ability, since her working memory is not yet as developed as the adult's (Reinhart, 1999). The eye-tracking research (Trueswell *et al.*, 1999; Snedeker *et al.*, 2001) shows that difficulty with utilizing referential information is due not to a general lack of knowledge that the discourse reference exists, but to memory limitations that prevent children from considering improbable syntactic alternatives. Only after the developing processing system gains the ability to maintain parallel parses over numerous words may contextual facts further drive processing decisions. And, indeed, the use of the eye-tracking methodology

in our Experiment 2 allowed us to uncover a major difference between children's off-line referent preference and their on-line ambiguity resolution. Contrary to the above-mentioned accounts, our results demonstrate that children unconsciously can access multiple referential representations.

A quite different pattern of children's awareness of sentence-external reference emerges if we compare the children's on-line eye movement data with those of adults'. Had we relied only on the picture selection task or any other traditional off-line methods such as act-out, we would have not discovered the striking discrepancy between how children resolve referential ambiguity off-line and on-line. While in the picture selection task the children overwhelmingly prefer the sentence-internal interpretation, their eye movements convincingly show their emerging awareness of the referential ambiguity, although they are not yet capable of expressing it consciously. Children's eye movement data were qualitatively similar to adults': the children showed the same increased looks to the sentence-external referent picture in the sentences with pronouns demonstrating competition between two possible referents. The children, however, were much slower than adults in noticing the ambiguity; they start to link the ambiguous pronoun with a potential antecedent on average 1000 ms later than adults, who, by then, have already resolved the referential ambiguity and selected a referent. This supports the hypothesis that children's processing mechanisms work just like adults'. What distinguishes sentence processing in children and adults is not a general inability to use referential information on the part of the children. Children need more time for accessing discourse reference and integrating discourse information during sentence processing than adults do.

Why, then, are the children unable to revise their initial commitment to the sentence-internal interpretation displaying deterministic rather than probabilistic parsing? If the eye movements are really reflective of the processing of choosing a referent for ambiguous pronouns, why is there this dissociation between action, e.g. a picture choice and making a decision in the case of children, but not adults? Note that Trueswell *et al.* (1999) found that children were unable to revise their initial commitment to an interpretation that turned out to be incorrect when later in the sentence the incompatible linguistic information was encountered. In our experiment, there was no incompatible information later in the sentence. The sentences were globally ambiguous, and both referents were potentially possible. We suggest that this dissociation between performance on a looking measure and a pointing measure for children in referential ambiguity resolution can be explained by a general pattern of difference between implicit knowledge and explicit understanding discussed in the recent literature on cognitive development. Recent findings in false belief tasks (Ruffman, Garnham, Import & Connolly, 2001) demonstrate that three-year-old children

sometimes look to the correct location, but bet very highly on the incorrect location consistent with their explicit incorrect verbal answer. This suggests that they were not aware of the knowledge conveyed through their eye gaze. Similarly, in the dimensional change card sort task, Zelaso, Frye & Rapus (1996) found a dissociation between preschoolers' awareness of the rules and ability to execute them. This dissociation has been attributed to a lack of development of executive functions of the prefrontal cortex responsible for self-regulation, inhibition, planning and modifying behaviour, as well as maintaining the representations in working memory (Robin & Holyoak, 1998). The discrepancy we have found between implicit resolution of referential ambiguity on-line and the deterministic choice of the sentence-internal referent by four-to-seven-year-old children brings in new evidence from a direct language task in favor of a general developmental dissociation between implicit knowledge and explicit understanding in the human processing system. Children's eye movements reveal implicit awareness of referential ambiguity that develops earlier than their explicit knowledge reflected in the picture selection task. As the human processing system matures and implicit awareness gradually shifts into explicit knowledge, deterministic parsing decreases and probabilistic parsing becomes increasingly common. Determining the precise developmental stage at which children begin to parse probabilistically remains a topic for future research.

REFERENCES

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S. & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eyetracking. *Cognition* **76**, B13–B26.

Arnold, J. E., Brown-Schmidt, S., Trueswell, J. & Fagnano, M. (in press). Children's use of gender and order of mention during pronoun comprehension. In J. C. Trueswell & M. K. Tanenhaus (eds), *Processing world-situated language: bridging the language-as-product and language-as-action traditions*. Cambridge: MIT Press.

Avrutin, S. (1994). *Psycholinguistic investigations in the theory of reference*. Unpublished doctoral dissertation, MIT.

Booth, J., MacWhinney, B. & Harasaki, Y. (2000). Developmental differences in visual and auditory processing of complex sentences. *Child Development* **71**, 981–1003.

Chien, Y.-C. & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* **1**(3), 225–95.

Chomsky, N. (1981). *Lectures in government and binding*. Dordrecht: Foris.

Crain, S. & Steedman, M. (1985). On not being led up the garden path: the use of context by the psychological parser. In D. Dowty, M. Karttunen & A. Zwicky (eds), *Natural language parsing*. Cambridge: CUP.

Crawley, R. A., Stevenson, R. J. & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research* **19**, 245–64.

Forster, K. & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers* **35**(1), 116–24.

Garrod, S. (1994). Resolving pronouns and other anaphoric devices: the case for diversity in discourse processing. In C. Clifton, L. Frazier & K. Rayner (eds), *Perspectives on sentence processing*. Hillsdale, NJ: Erlbaum.

Garvey, C., Caramazza, A. & Yates, J. (1974/5). Factors influencing assignment of pronoun antecedents. *Cognition* **3**, 227–43.
Gerken, L. & Shady, M. E. (1996). The picture selection task. In D. McDaniel, C. McKee & H. S. Cairns (eds), *Methods for assessing children's syntax*. Cambridge: MIT Press.
Gernsbacher, M. A. & Hargreaves, D. J. (1988). Accessing sentence participants: the advantage of first mention. *Journal of Memory and Language* **27**, 699–717.
Goodluck, H., Terzi, A. & Chocano Díaz, G. (2001). The acquisition of control cross-linguistically: structural and lexical factors in learning to licence PRO. *Journal of Child Language* **28**, 153–72.
Gordon, P. C. & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition* **23**, 313–23.
Grodzinsky, Y. & Reinhart, T. (1993). The innateness of binding and coreference: a reply to Grimshaw and Rosen. *Linguistic Inquiry* **24**, 69–101.
Hahne, A. & Friederici, A. (2001). Development patterns of brain activity. In J. Weissenborn & B. Höhle (eds), *Approaches to bootstrapping: phonological, lexical, syntactic and neurophysiological aspects of early language acquisition*. Vol. 2. Amsterdam: John Benjamins.
Holcomb, P. J., Coffey, S. A. & Neville, H. J. (1992). Visual and auditory sentence processing: a developmental analysis using event-related brain potentials. *Developmental Neuropsychology* **8**(2–3), 203–41.
McDaniel, D., McKee, C. & Cairns, H. S. (eds) (1996). *Methods for assessing children's syntax*. Cambridge: MIT Press.
McKee, C. (1996). On-line methods. In D. McDaniel, C. McKee & H. S. Cairns (eds), *Methods for assessing children's syntax*. Cambridge: MIT Press.
Maat, H. P. & Sanders, E. (2001). Subject–pronoun competition in resolving ambiguous anaphors. Paper presented at the workshop 'From Sentence Processing to Discourse Interpretation: Crossing the Borders'. Utrecht University.
Musolino, J., Crain, S. & Thornton, R. (2000). Navigating negative quantificational space. *Linguistics* **38**, 1–32.
Nadig, A. & Sedivy, J. (2000). Children's use of referential pragmatic constraints in production and processing. Paper presented at the 13th Annual CUNY Conference on Human Sentence Processing.
Piñango, M. M., Burkhardt, P., Brun, D. & Avrutin, S. (2001). The psychological reality of the syntax-discourse interface: the case of pronominals. Paper presented at the workshop 'From Sentence Processing to Discourse Interpretation: Crossing the Borders'. Utrecht University.
Reinhart, T. (1999). *The processing cost of reference-set computation: guess patterns in acquisition*. OTS working papers (Uil-Ots 99001-CL/TL). Utrecht University.
Robin, N. & Holyoak, K. J. (1998). Relational complexity and the functions of prefrontal cortex. In M. S. Gazzaniga, R. B. Ivry & G. R. Morgan (eds), *Cognitive neuroscience: the biology of mind*. New York: WW Norton & Company.
Ruffman, T., Garnham, W., Import, A. & Connoly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology* **80**, 201–24.
Runner, J. T., Sussman, R. S. & Tanenhaus, M. K. (in press). Assignment of reference to relatives and pronouns in picture noun phrases: evidence from eye movements. *Cognition*.
Smyth, R. H. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* **23**, 197–229.
Snedeker, J., Thorpe, K. & Trueswell, J. (2001). On choosing the parse with the scene: the role of visual context and verb bias in ambiguity resolution. *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
Swinney, D. & Prather, P. (1989). On the comprehension of lexical ambiguity by young children: investigations into the development of modularity. In D. S. Gorfein (ed.), *Resolving Semantic Ambiguity*. Berlin: Springler-Verlag.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* **268**, 1632–34.

Tanenhaus, M. & Trueswell, J. (1995). Sentence comprehension. In J. Miller & P. Eimas (eds), *Speech, language, and communication*. San Diego: Academic Press.

Tenny, C. L. (1999). *Evidentiality and the grammaticalization of sentience. Three phenomena involving evidentiality*. Talk presented at the University of Rochester.

Thornton, R. & Wexler, K. (1999). *Principle B, VP ellipsis, and interpretation in child grammar*. Cambridge: MIT Press.

Traxler, M. (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: evidence from self-paced reading. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* **55**A(1), 75–96.

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

Tyler, L. K. & Marslen-Wilson, W. (1981). Children's processing of spoken language. *Journal of Verbal Learning and Verbal Behavior* **20**, 400–16.

Wykes, T. (1981). Inference and children's comprehension of pronouns. *Journal of Experimental Child Psychology* **32**, 264–78.

Zelaso, P. D., Frye, D. & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development* **11**, 37–63.