How do adults and children process referentially ambiguous pronouns?

IRINA SEKERINA
City University of New York
KARIN STROMSWOLD
Rutgers University
AND
ARILD HESTVIK
University of Bergen and City University of New York

ABSTRACT

A recently developed method of head-mounted eye-tracking (Trueswell et al., 1999) has been employed in two experiments to investigate differences between adults and children in on-line processing of ambiguous short-distance pronouns (Tenny, 1999) in English. Sixteen adults and 16 five-to-seven-year-old children viewed a pair of pictures with two characters and listened to sentences describing the pictures which contained either an unambiguous reflexive (himself) or an ambiguous pronoun (him). They had to choose a picture that corresponds to the sentence. For adults, all three types of analyzed data — responses, reaction times, and eye movements — indicate that pronouns are referentially ambiguous (20% sentence-external referent). Adults were capable of making inferences on-line with respect to less accessible referent, showing a competition between the looks to two potential referents shortly after encountering the pronoun in the sentence. Children’s response data reveal that they overwhelmingly prefer the sentence-internal referent for the ambiguous pronoun (93%). However, their eye movement data are qualitatively similar to adults’, showing their emerging adult-like awareness of the referential ambiguity in sentences with pronouns. But since children’s processing resources are more limited than those of adults (Chien and Wexler, 1990; Avrutin, 1994), children need more time for accessing the sentence-external referent and are reluctant to revise their initial sentence-internal referent commitment in the absence of information that would require such a revision in globally ambiguous sentences. This supports the hypothesis that children’s processing mechanism works like adults’ (Thornton and Wexler, 1999).

1. Introduction

The acquisition of the two classes of pronominals, pronouns such as him and her and reflexives such as himself and herself in English 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, or by interpretative rules operating at the LF-discourse interface, depending on the environment.1 If a pronoun is syntactically bound, its reference is determined by the Binding Theory and the A-Chain Condition (Reinhart and Reuland, 1993). According to the Binding Theory, reflexives obey Principle A (Chomsky, 1981 and thereafter) while pronouns obey Principle B. In a sentence such as (1), the reflexive himself is interpreted as coreferential with the subject of the clause, the boy: According to Principle A, it is bound by an antecedent in its governing category, the tensed clause. In contrast, the pronoun him in (2) must be free in its governing category, as required by Principle B.

(1) The boy$_1$ has washed himself$_1$.
(2) The boy$_1$ has washed him$_{2\star 1}$.

The basic intuition derived from the contrast in (1) and (2) is that reflexives and pronouns are in complementary distribution. The linguistic literature, however, sites a number of cases where this complementarity breaks down; these are classical exceptions to the principles of the Binding Theory. An example involves pronouns that take an antecedent within a local domain where Principle B should rule them out, as illustrated by (3). Assuming the local domain to be the clause containing the pronoun inside the prepositional phrase (PP) in (3b), we predict that him cannot refer to the boy. Surprisingly, the sentence is referentially ambiguous: the pronoun can be coreferential either with a sentence-external referent, a

1 Alternative linguistic accounts of referential coherence include Ariel, 1990; Givón 1992; Walker, Joshi, and Prince, 1998.
difficulty in assigning further occurs in the sentence itself rather than the surrounding context. Thus, one may imagine should be more accessible than

Tenny (1999) refers to this phenomenon as Short-Distance Pronouns (SDP) and identifies three productive environments for SDPs in English: locative or directional prepositional phrases (3b), colloquial sentences (4a), and representational noun phrases (4b). In this paper, we will be concerned with the locative PPs only.

The sentence-external interpretation of the pronoun in the directional prepositional phrases in (3b) is not regulated by purely syntactic constraints; rather, it is constrained by the Interpretation Rule, Rule I, at the LF-Discourse interface (Grodzinsky and Reinhart, 1993). Since the subject NP the boy and the pronoun him are allowed to be counterindexed, Rule I causes counterindexation to entail noncoreference.

The questions of why the SDPs bound by the sentence-internal referent are exempt from Principle B, and how to account for this phenomenon from the point of view of linguistic theory and sentence processing, are discussed in detail in the literature (Hestvik, 1992; Tenny, 1999; Runner, Sussman, and Tanenhaus, 2001). The sentence-external referent, on the other hand, does not present a problem from the syntactic point of view since it is regulated by Rule I. Rule I is one of the heuristics for the interpretation of referential expressions in terms of the preceding discourse (Garrod, 1994; Gordon and Searce, 1995). More specifically, current linguistic and processing theories contain assumptions about the preferred order in which possible interpretations will be considered during processing. It is predicted that the most recent referent, the subject of the clause, is a more prominent candidate than referents appearing in other grammatical roles; that is, the subject is more accessible (Garrod, Freudenthal, and Boyle, 1994; Arnold, 1998; Arnold, Eisenband, Brown-Schmidt, and Trueswell, 2000). In (3b), for example, the NP the boy should be more accessible than the man because the boy is the grammatical subject of the clause and furthermore occurs in the sentence itself rather than the surrounding context. Thus, one may imagine difficulty in assigning him to the sentence-external referent the man since the Subject rule will have to be overruled: the antecedent is only introduced by the discourse and is not mentioned in the sentence. Interestingly, while the linguistic theory allows SDPs to freely choose the sentence-internal or –external interpretation, the psycholinguistic expectations are that the Subject rule and the accessibility and first-mention factors should favor the sentence-internal referent for ambiguous pronouns. Only stronger constraints will be able to outweigh them in order to enhance the prominence of the sentence-external referent for SDPs. For example, when the availability of the sentence-external referent, e.g., the man, is promoted by the persuasive context, its discourse prominence makes it a valid competitor against the sentence-internal referent the boy. Surprisingly, there are no experimental studies in adult psycholinguistic research that either support or refute the linguistic assumption of referential ambiguity and test the hypothesized processing preferences.

Consequently, we do not know about children’s knowledge of SDPs and how they resolve this referential ambiguity. Numerous language acquisition studies have shown that children often have difficulty using Rule I no matter which language and what construction (Avrutin, 1994; Chien and Wexler, 1990; Grodzinsky and Reinhart, 1993; Hestvik and Philips, 1999/2000), which should make the sentence-external interpretation difficult to access. Moreover, young children overgeneralize coreferential readings with counterindexed NPs, allowing him to take the boy as an antecedent in much simpler sentences such as (2). It would not be surprising if children were to undergeneralize in (3b) by exclusively selecting the sentence-internal referent, although this preference may potentially be weakened by the fact that children’s use of recency and order-of-mention is not completely adult-like (Arnold et al., 2000). One of the explanations proposed to account for difficulties children experience with sentence-external referents relies on the hypothesis that children’s processing and inferential resources are limited (Paris and Lindauer, 1976; Wykes, 1981; Grimshaw and Rosen, 1990; Goodluck, 1990; Chien and Wexler, 1990; Avrutin, 1994; Reinhart, 1999; Trueswell et al., 1999, among others). While the properties of the adult processing system have been in the focus of experimental psycholinguistics for the past 30 years, we are just now beginning to discover how children develop processing mechanisms. The important issue at hand is whether adults and
children have access to the same universal parsing routines or not. It is notoriously difficult to study children’s language comprehension on-line due to lack of child-friendly experimental techniques (cf. Booth, MacWhinney, and Harasaki, 2000; Hahne and Friederici, 2001; Trexler, 2001). Recently, with the advance of the free head-mounted eye-tracking methodology, researchers started to investigate the development of processing mechanisms in children (Trueswell, Sekerina, Hill, and Logrip, 1999). Several other eye-tracking studies of how children process language on-line followed, investigating various aspects of children’s processing ranging from grammatical gender (Arnold et al., 2000) and contrastive focus (Nadig and Sedivy, 2000) to verb lexical bias (Snedeker, Thorpe, and Trueswell, 2001). But neither of them has addressed the question of how children deal with pronoun resolution in real-time. Given this gap, we investigated how four-to-seven-year-old English-speaking children process referentially ambiguous pronouns in a moment-by-moment fashion. We compared children’s and adults’ referent preference data using both off-line and on-line methods in order to empirically test the qualitative differences between the child’s processing system and the adult parser.

This article is organized as follows: Section 2 summarizes recent findings about the factors that contribute to referential ambiguity resolution in adult sentence processing. In Section 3, we present the results from an off-line referent preference questionnaire that establishes the baseline for adults in choosing a possible referent for the pronoun in the locative PPs. Then, using the technique of monitoring adult participants’ eye movements while they listen to spoken sentences (Tanenhaus, Spivey-Knowlton, Eberhard and Sedivy, 1996; Allman and Kamide, 1999; Arnold et al., 2000), we investigate how the established referential ambiguity is resolved on-line. This method allows us to monitor how listeners look for possible referents for pronouns in a real-time fashion without interrupting the comprehension process. In Section 4, we report the results of the children’s on-line eye-tracking study and compare them with the collected adult data. Since this method provides moment-by-moment information about the referents that children are considering for both ambiguous and unambiguous pronouns, it gave us an opportunity to discover the striking discrepancy between how children resolve referential ambiguity off-line and on-line. While the behavioral data indicate that children overwhelmingly prefer the sentence-internal referent, their eye movements convincingly show that they are aware of the referential ambiguity although they are not yet capable of expressing it consciously. Finally, in Section 5, we summarize our findings and discuss their significance for adult sentence processing theory and for research on development of processing mechanisms.

2. Factors that affect referential ambiguity resolution for pronouns

Accessibility, order-of-mention, and recency have all been proposed as important discourse-related factors that are used to guide referential processing in the large body of literature on pronoun resolution (Carreiras, Garnham, and Oakhill, 1993; Corbett and Chang, 1983; Crawley, Stevenson, and Kleinman, 1990; Garvey, Caramazza, and Yates, 1974/5; Gernsbacher and Hargreaves, 1988; Smyth, 1994). Several recent psycholinguistic studies have investigated referential ambiguity resolution for ambiguous pronouns in order to examine the psychological reality of syntax-discourse relations. Piñango, Burkhardt, Brun and Avrutin (2001) examined sentences where the competing referent is arbitrary (5b):

(5) a. Everyone1 thinks that students like him1,…
     b. The teacher1 thinks that students like him1,…

The reaction times in the lexical decision task for the referring pronouns in (5b) were significantly higher than for the bound variable pronouns in (5a). The researchers argued for the discourse processing complexity hypothesis: Processing of (5a), where interpretation for the bound variable pronoun him is obtained via syntactic mechanisms alone, is less costly than that of (5b). In the latter case, the language processor has to access information beyond that provided in the syntax; that is, information from discourse.

Maat and Sanders (2001) conducted a corpus study in Dutch with potentially ambiguous newspaper fragments, in which a sentence containing the pronoun he (hij in Dutch) was preceded by a sentence with two male singular participants, as in their English counterparts in (6):

(6) a. George1 hit Al2. He1/2 was really mad.
     b. George1 hit him2. He1/2 was really mad.
The results showed that the Subject rule indeed appeared to dominate pronoun resolution in (6a): in 77% of the cases, the actual referent occurred in subject position, i.e., *he* referred to *George*. However, the results were different in (6b), where the competing referent was a pronoun itself: *he* was the referent for *him* in 68% of the cases. Maat and Sanders proposed the Pronoun rule to account for this preference and argued that referential ambiguity resolution is regulated by interaction of the Subject and Pronoun rules, along the lines of the constraint-satisfaction approach.

Arnold and colleagues (Arnold et al., 2000) took this investigation one step further and examined moment-by-moment resolution of referential ambiguity in English sentences similar to (6a), in which both potential referents were lexical NPs. The materials consisted of four types of pictures representing two characters who were either both males (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:

(7) a. Donald is bringing some mail to Mickey while a violent storm is beginning.
   b. He’s carrying an umbrella, and it looks like they’re both going to need it.

(Arnold et al., 2000: B18 (3))

The head-mounted eye-tracking system was used to record subjects’ eye movements as they looked at the pictures and listened to the sentences in (7). In the example with the gender ambiguity, Arnold and colleagues found that subjects did not immediately converge on an interpretation of the pronoun referent (Mickey); instead, they looked equally at both characters for a while. Mickey was the only correct referent because the picture depicted him as the one carrying the umbrella. 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 (7a), *Donald is bringing some mail to Mickey. He’s sauntering down the hill, while a violent storm is beginning.* In this case, pronominal reference to the second character (Mickey) was highly infelicitous, and subjects often made mistakes in these sentences. In contrast to the first experiment, participants’ eye movements revealed no competition between Donald and Mickey, providing evidence for order-of-mention and recency as sources of information that guide on-line resolution of referential ambiguity.

The two adult studies that follow contribute to this ongoing research on how adults and children look for referents for ambiguous pronouns. Experiment 1, the off-line referent preference questionnaire, establishes the baseline for choice of the referent and empirically tests the linguistic hypothesis about referential ambiguity of SDPs. Experiment 2, the 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.

3. Processing of referentially ambiguous pronouns by adults

3.1. Experiment 1: Referent preference questionnaire

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.

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, in which a questionnaire was completed requiring judgments about interpretation preferences of pictures. Completion of the questionnaire typically took 20 minutes.
Design and materials. Participants were presented with 24 pairs of pictures, illustrated in Figure 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. For example, in Fig. 1(A), the two characters were a boy and a man, and the object was a box. The two pictures were identical except in the right picture the boy is wearing a cap, and in Picture A, the box is behind the boy, and in Picture B, it is behind the man. The pairs of pictures always contrasted either two female characters (a girl, a woman, or a grandmother), as in Fig. 1(B), or two male characters (a boy, a man, or a grandfather) (Fig. 1A). 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. In total, 12 different objects were used: bucket, box, hairbrush, horse, teddy bear, pig, glass, hat, suitcase, ball, chair, and watering can, with each object appearing 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 distributed in a counterbalanced design.

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).

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 (REFL condition) or with the pronoun (PRON condition). The questionnaire was assembled according to a design schematically represented in Fig. 2. (The full set of the experimental materials can be found in the Appendix.)

24 experimental questions

<table>
<thead>
<tr>
<th>12 REFL condition</th>
<th>12 PRON condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 FEMALE</td>
<td>6 MALE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3 next</th>
<th>3 behind</th>
<th>3 next</th>
<th>3 behind</th>
<th>3 next</th>
<th>3 behind</th>
<th>3 next</th>
<th>3 behind</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 put</td>
<td>3 placed</td>
<td>3 put</td>
<td>3 placed</td>
<td>3 put</td>
<td>3 placed</td>
<td>3 put</td>
<td>3 placed</td>
</tr>
</tbody>
</table>

Fig. 2. The factors manipulated in the experimental stimulus questions

So that identical materials would not be repeated to any subject, the questionnaire was constructed in two versions, with the REFL and PRON conditions items distributed over versions in a counterbalanced design. There were no filler items in the questionnaire. Thus, each version of the experiment contained 24 picture-

---

2 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 and Philip (1999/2000). The authors would like to Ned Norland for his assistance with modification of the original pictures.
pairs-plus-text items in total, with the items in the two conditions interspersed in a pseudo-random order. The picture-pair-plus-text item consisted of two preamble sentences, an experimental question, and a set of three possible answers, as illustrated in (8).

(8) PREAMBLE: In these pictures, you see a boy, a man, and a box. The boy has placed the box on the ground.

QUESTION:
REFLEXIVE CONDITION: Which picture shows that the boy has placed the box behind himself?
PRONOUN CONDITION: 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, as highly accessible in spoken discourse. As example (8) demonstrates, it was mentioned three times, twice in the preamble and once in the experimental question. Thus, the boy was easily accessible due to its salience and recency. In contrast, the sentence-external referent, the man, was not as accessible as the sentence-internal referent because the former was mentioned only once and was separated from the pronoun by two sentences. Visually, however, both referents were equally salient.

Procedure and Data Treatment. 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, to read the preamble sentences and the question carefully, and then to indicate their answer to the question by circling one of the three possible answers. The responses in the REFL condition were screened for errors, and two participants with 50% or more errors (six out of 12 items) were rejected. For purposes of analysis, the response data were assembled into 3 x 2 x 2 x 28 matrices. This analysis design brings together the referent preference (made by 28 participants in each of the two versions of the experiment) for three exemplars of the two prepositions and two genders. Two such matrices were prepared, one for each of the conditions, REFL vs. PRON. The data were analyzed parametrically in analyses of variance.

The terms of the ANOVA bear directly on the question around which the experiment was designed. The main effect of the Condition 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 REFL condition, presumably reflecting the ambiguous status of the pronoun in the PRON condition. The non-significance in the Preposition x Verb x Gender x Condition interaction term of the ANOVA indicates a reasonable consistency across questions in selecting the sentence-internal referent interpretation. This means that idiosyncratic 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 REFL questions were overall interpreted correctly: the participants chose the sentence-internal referent in 96% cases (20 participants out of 28 did so in 100% cases). That is, in the version of Item 27 with the reflexive (Fig. 1 and Example (8)), they chose Answer (a) “the left picture” indicating that himself takes the sentence-internal referent the boy as its antecedent. The responses in the PRON questions reflect the ambiguous referential status of the pronoun; overall, in approximately 20% of cases, the participants chose the subject-external referent for the ambiguous pronoun. That is, in the version of Item 27 with the pronoun, participants sometimes 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.
Table 1. Adult referent preference questionnaire: Overall sentence-internal referent preference

<table>
<thead>
<tr>
<th>Referent</th>
<th>Percentage choice for the pronominal</th>
<th>Sentence-internal</th>
<th>Sentence-External</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFL</td>
<td>95</td>
<td>5a</td>
<td>0a</td>
<td></td>
</tr>
<tr>
<td>PRON</td>
<td>79</td>
<td>17</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

*a Incorrect response.

The difference in amount of sentence-internal referent for REFL vs. PRON was highly significant (F(1,57)=24.11, F(2,23)=45.16, p<.0005). Figures in Table 2 demonstrate that apart from the Condition main effect, there were no significant main effects for any of the three additional lexical factors: verb, gender, or preposition.

Table 2. Referent preference questionnaire

<table>
<thead>
<tr>
<th>Referent</th>
<th>Male gender</th>
<th>Female gender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Behind</td>
<td>Near</td>
</tr>
<tr>
<td></td>
<td>Put</td>
<td>Placed</td>
</tr>
<tr>
<td>REFL</td>
<td>96</td>
<td>97</td>
</tr>
<tr>
<td>PRON</td>
<td>83</td>
<td>80</td>
</tr>
</tbody>
</table>

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 still chose the sentence-external interpretation for the pronoun in 20% 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 (between (7a) and (7b) above) made the latter highly infelicitous and resulted in listeners choosing an incorrect but more accessible referent. Thus, adult English speakers in Experiment 1 indeed found the sentences with SDPs ambiguous, supporting the claim made in the linguistic literature.

3.2. Experiment 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 based on the accessibility of referent information. How do participants make this choice? The off-line nature of Experiment 1 did not provide any information about how the referential ambiguity is considered and resolved. Our 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, in which they were required to choose the picture representing their reference preference by pressing a corresponding button. Completion of the experiment typically took 40 minutes.

Design and materials. The materials used in Experiment 2 included the 24 experimental items from Experiment 1 and 30 fillers. The fillers also always contained a pair of pictures, sometimes very similar to each other (Fig. 3A), so as to resemble the experimental trials, and sometimes quite different (Fig. 3B).
The instructions in this experiment were longer than those in Experiment 1. They consisted of four or five sentences: one preamble sentence, the fixation instruction, the experimental question, and one or two follow-up instructions, as illustrated in (9) for experimental item 27 (Fig.1A) and in (10) for filler 2 (Fig. 3B).

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

Note that the accessibility of the two potentials referents was still different, although the sentence-external referent, the man, was potentially more accessible in (9) than in (8). Only one sentence separated the pronoun and the referent, and this sentence did not contribute to the prepositional situation. The sentence-internal referent, the boy, remained highly accessible.

(10) 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.

The same experimental manipulations from Experiment 1 (i.e., gender, verb type, and preposition type) were used in Experiment 2. Two versions of the experiment were prepared in a counterbalanced design, with 54 items in each.

Procedure and Data Treatment. 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 (see Fig. 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 stimuli 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). The cross-hair was not projected on the stimuli computer screen and thus was 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 about 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 stimuli 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).\(^3\) 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. Once the eye image was judged to be suitable, the participant had to hold his or her gaze on a series of five spatial positions 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” Dell desktop monitor, under the control of DMDX software (Forster and Foster, 1990). Each display consisted of two pictures of equal size separated by a line with a cross-hair in the middle (see Fig. 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, 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., \textit{In these pictures, you see a boy, a man, and a box}) and a request to look at the central fixation point (\textit{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 phonetician\(^4\) 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”. The response data in the REFL condition were screened for errors, and since every participant produced highly accurate results in this condition, no one was rejected according to this criterion. In addition, the eye-

\(^3\) We refer the reader who is interested in a more detailed description of the ISCAN eye-tracking system to Trueswell, Sekerina, Hill, and Logrip (1999, Section 4.1: 98-100).

\(^4\) We thank Jennifer Venditti for providing her expertise in analyzing the materials with the help of the ToBI system (Beckman and Elam, 1994).
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 analyzed in this experiment: responses in the forced-choice task, reaction times for the forced-choice task, and eye movements. First, just as in Experiment 1, the response data were assembled into 3 x 2 x 2 x 16 matrices that brought together the referent preference (made by 16 participants in each of the two versions of the experiment) for three exemplars of the two prepositions and two genders. Two such matrices were prepared, one for each of the conditions, REF vs. PRON. The data were analyzed parametrically in analyses of variance. Second, the reaction time data were analyzed for information about whether participants took longer to choose a picture in the PRON condition compared to the REF 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. Statistical significance in this analysis indicates that longer reaction times in the PRON 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 during the course of interpreting spoken sentences, and provided information about 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 response data obtained in Experiment 2 closely resemble those from Experiment 1 (Table 3).

<table>
<thead>
<tr>
<th>Percentage referent choice for the pronominal</th>
<th>Sentence-internal</th>
<th>Sentence-External</th>
</tr>
</thead>
<tbody>
<tr>
<td>REF</td>
<td>97</td>
<td>3^a</td>
</tr>
<tr>
<td>PRON</td>
<td>80</td>
<td>20</td>
</tr>
</tbody>
</table>

^a Incorrect response.

Participants chose statistically significant higher percent of the sentence-internal interpretation for REF than for PRON (F1(1,15)=4.63, p<.05, F2(1,23)=46.79, p<.0005). There were no significant main effects for gender, verb, or preposition. Thus, the on-line results of Experiment 2 replicated those of Experiment 2: the PRON sentences are ambiguous for adults, although they do prefer sentence-internal referents.

Reaction times. Table 4 below summarizes the reaction time data, for each of the two conditions. As Table 4 shows, the reaction time data were split in two ways. First, all responses — 97% sentence-internal responses in the REF condition vs. 80% sentence-internal and 20% sentence-external responses in the PRON condition; second, only for the responses in the PRON condition — 80% sentence-internal vs. 20% sentence-external responses.

<table>
<thead>
<tr>
<th>Reaction times (ms)</th>
<th>REF</th>
<th>PRON All</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-Internal</td>
<td>1245</td>
<td></td>
</tr>
<tr>
<td>S-External</td>
<td>1573</td>
<td></td>
</tr>
</tbody>
</table>

5 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.

6 Since the task was forced-choice, only two types of responses were possible: the sentence-internal interpretation and the sentence-internal interpretation. In Experiment 1, the additional option was “both”.

---

10
As the reaction times indicate, participants were faster for REFL than for PRON, although this difference approached significance in the item-based analysis only (F1(1,15)=2.49, p=.136, F2(1,23)=3.22, p=.086). This difference becomes substantially larger when the reaction times for the pronoun sentences are split into RTs for the sentence-internal vs. the sentence-external referent preference (Δ328 ms).

Eye movements. Prior to the eye movement analysis we calculated the onsets of the three regions in the sentences (Fig. 5). Region 1 started at the onset of which in Which picture shows that… So that all experimental items would share the same start time, the beginning of Region 1 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. The verb placed was consistently longer than the verb put, as was the preposition behind compared to near. 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).

<table>
<thead>
<tr>
<th>Which picture shows that</th>
<th>the boy has placed the box behind</th>
<th>him himself?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region 1</td>
<td>Region 2</td>
<td>Region 3</td>
</tr>
<tr>
<td>0</td>
<td>1000 ms</td>
<td>2475 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4475 ms</td>
</tr>
</tbody>
</table>

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

Before we proceed to discussion of the results, we need to address the issue of missing data. Recall that subjects’ task was to perform speeded force-choice task: they pressed the right or left button to indicate which of the two pictures they chose. Theoretically, we would expect the majority of button presses to fall into Region 3 since the subjects presumably could not make a meaningful choice before they heard and processed the entire sentence. However, as is typically the case with the reaction time data, in a certain number of trials, the subjects pushed the button ahead of time, without waiting until the end of the sentence. Once the button was pushed, the trial was considered completed and recording of eye movements stopped. In our Experiment 2, we had the following percentages of trials completed in each of the three regions: Region 1 (0 -1000 ms): 16%, Region 2 (1000-2474 ms): 17%, and Region 3 (2475-4475 ms): 63%. Thus, the eye movements results 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.

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 on which of the two pictures the participant was fixating. Thus, for Item 27, The boy has placed the box

---

7 Due to the fact that the sentence-external responses were distributed unevenly across the subjects and items, it was impossible to carry out the analysis of variance on the sentences with pronouns only.

8 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).

9 Thus, approximately 4% of the trials were completed after the selected stop point.

10 We wish to thank Dr. Gary Winkel, professor of Psychology at the Department of Psychology at the CUNY Graduate Center for providing statistical expertise and help with the statistical analyses reported in this article.
behind him (see Fig. 1 above), the looks to the left picture indicate the sentence-internal interpretation and to the right one the sentence-external interpretation.11

Fig. 6 presents the graphical results of the coarse-grain analysis. It shows the percentage of fixations of either the sentence-internal or sentence-external picture for each 33-ms frame in the entire Region 3, starting after the onset of him/himself (75 frames=2475 ms). There were more looks to the sentence-internal picture in the REFL condition than in the PRON condition. In contrast, there were more looks to the sentence-external picture in the PRON condition than in the REFL condition. This indicates that participants were aware of the referential ambiguity in the PRON condition and were resolving it on-line.

![Fig. 6. Adult Eye-Tracking Experiment 2: Probability of fixating the Sentence-Internal referent picture in two conditions (%), for the entire Region 3](image)

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 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 Fig. 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 REFL condition than in the PRON condition, 65% vs. 55% (F(1,15)=5.71, p=.0304). The same pattern was found for Segment 1, main effect of Condition, 63% vs. 55% (F1(1,15)=4.83, p=.0441), and Segment 2, 69% vs. 55% (F(1,15)=4.85, p=.0463).12 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

---

11 In addition to the coarse-grain analysis, we also performed fine-grained analysis for which we coded the looks within each of the two pictures for three separate locations, Referent, Object, and Other. Three additional fixation locations were also coded: Cross (the central fixation point), Else (looks elsewhere in the picture), and Track Loss. Together, these three accounted for 15% of all the eye movement data (Cross 10%, Else .5%, and Track Loss 4.5%). The results of the fine-grained analysis are not reported here since they do not directly bear on the question of how the referential ambiguity is resolved on-line.

12 There was also a marginal effect of Preposition, 65% near vs. 54% behind (F1 non significant, F2(1,22)=4.037, p=.057.).
rapidly took into consideration the alternative possible interpretations. The resolution of the referential ambiguity took place soon after the lexical information about the pronoun became available; around 250-350 ms after the onset of the pronoun **him/himself**, adults engaged in anaphoric resolution, and resolved the referential ambiguity on-line. This time course is comparable to rapid and incremental referential interpretation of lexical NPs (Tanenhaus *et al.*, 1995; Trueswell *et al.*, 1999; Altmann and Kamide, 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 analyzed 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. Since adults’ processing resources are not as limited as children’s, they are 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 (Paris and Lindauer, 1976; Wykes, 1981; Grimshaw and Rosen, 1990; Chien and Wexler; Avrutin, 1994; Trueswell *et al.*, 1999), they will fail to resolve this referential ambiguity? We address this question in a children’s eye-tracking experiment.

### 4. Processing of referentially ambiguous pronouns by children

#### 4.1. Acquisition of Pronouns and Development of Processing Mechanism

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 and Wexler, 1990; Grimshaw and Rosen, 1990; Grodzinsky and Reinhart, 1993; Kaufman, 1994). However, short-distance pronouns, the focus of the present article, 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, that is, to take either a sentence-internal or sentence-external referent, poses familiar problems of learnability (Musolino, Crain, and Thornton, 2000). Theoretically, the choice of a sentence-internal or sentence-external referent may be optional, obligatory, or impossible depending on the lexical nature of the pronominal element involved (reflexive vs. pronoun and its position in the sentence.) If the child undergeneralizes possible referential interpretations, she may assume that a sentence containing a pronoun in the SDP context could only receive the sentence-internal interpretation, when that sentence may in fact allow both interpretations in the adult grammar, as we have demonstrated in the two adult experiments above. The acquisition literature often ascribes this undergeneralization to the fact that pre-school children have not yet mastered certain discourse mechanisms, including the Referential Principle (Crain and Steedman, 1985), Principle P (Chien and Wexler, 1990) or Rule I (Grodzinsky and Reinhart, 1993). Ultimately, children reach adults’ knowledge of binding principles, leading to the assumption that this knowledge is encoded in children’s grammar from the earliest stages of language development. These observations motivated our Experiment 3, 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 parsers.

Thornton and 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 know about how children process language in real-time due to the limitations in processing techniques for studying children’s language comprehension (Holcomb, Coffey, and Neville, 1992; McKee, Nicol, and McDaniel, 1993; Swinney and Prather, 1989; Tyler, 1983; Tyler and Marslen-Wilson, 1981). This issue only now becoming the focus of empirical and systematic investigation. Booth, MacWhinney, and Harasaki (2000)

---

13 For an overview, see Section 2 *Children’s on-line language processing abilities* in Trueswell *et all.* (1999).
and Traxler (2001) 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 and 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. The pioneering research by Trueswell et al. (1999) was the first to examine whether much younger, pre-literate children of four to six years of age employ syntactic and referential factors on-line in order to resolve syntactic ambiguity. This study and several others successfully employed the head-mounted eye-tracking system to monitor children’s eye movements while the children listened to the spoken instructions and thus did not rely on children’s yet to be developed reading skills (Arnold et al., 2000; Nadig and Sedivy, 2000; Snedeker, Thorpe, and Trueswell, 2001). Since these previous studies established that the head-mounted eye-tracking technology can successfully tap into children’s real-time processing, we employed it to study how children process referentially ambiguous pronouns in our Experiment 3.

If we relied only on the behavioral response data, it would appear that childrens’ grammar undergeneralizes SDPs. As we report below, the response data indicate that four- to seven-year-old children fail to access the sentence-external interpretation in the sentences with pronouns in SDP contexts. Instead, the children access only the sentence-internal interpretation taking the subject of the clause as the antecedent of the pronoun. This is in contrast to what we have found in Experiments 1 and 2: In the adult grammar, both interpretations are available, with a preference for the sentence-internal one. But a different pattern emerges when we examine children’s eye movement data. Children’s emergent awareness of the ambiguity reveals itself as they do look less at the sentence-internal picture in the sentences with pronouns than in the sentences with reflexives, a pattern found in adults’ eye movements in Experiment 2. However, these looks occur much later in the sentence, and children fail to pursue the sentence-external referent demonstrating once again (cf. Trueswell et al., 1999) their reluctance to revise the strong initial commitment to the sentence-internal interpretation.

4.2. Experiment 3: Children’s Eye-Tracking Experiment

Participants. The participants were sixteen children ranging in age from 4;9 to 7;10 (mean age 6;6, eight between 4;9 and 6;5 and eight between 6;9 and 7;10). As with the adults in Experiment 2, eight children were assigned to one version of the experiment and eight to the other. The completion of the experiment typically took 20 minutes. Children received a toy for their participation in the study.14

Design and materials. The target materials and design were basically the same as in Experiment 2, with several important modifications. First, the spoken instruction sentences were recorded, in contrast to the same instructions being produced live in Experiment 2. Second, 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 (see Appendix for the exact list of the children’s materials). The instructions were similar to Experiment 2 but consisted of only four sentences: one preamblesentence, the fixation instruction, the experimental question, and one follow-up instruction. The child’s task was also a forced-choice of one of the two pictures. However, the children did not have to press buttons; instead, they had to choose a picture by pointing to it with their finger. The instructions were adapted to accommodate pointing instead of looking, as illustrated in (11; cf. (8)) for experimental item 27 (see Fig. 1A).

(11) PREAMBLE: In these pictures, you see a boy, a man, and a box.
FIXATION: Now look at the cross.
QUESTION: 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 3. As before, the accessibility of the referents for the pronoun was kept consistent.

14 We wish to thank the students of Irving Elementary School (Speech coordinator Ellyn Atherton) and the children of the Yellow Brick Road Pre-school (Director Grace Puleo) in Highland Park, NJ, for their help in running the study.
across the conditions. Two versions of the experiment were prepared in a counterbalanced design, with 32 items in each.

**Procedure and Data Treatment.** While the adult Experiment 2 used a remote tabletop ISCAN eye-tracking system (see Section 3), an ISCAN head-mounted eye-tracking visor was used for the children. There were two reasons to switch to the head-mounted version of the eye-tracker. First, recall that for the remote eye camera, movements of the participant’s head were restricted to the range of 12 inches horizontally and four inches vertically. Legitimate head movements resulted in the automatic repositioning of the tilting pan, and the camera followed the eye within this range. When we piloted the remote system with two children, we found it impossible to restrict the children’s body and head movements within the required range. Second, the child participants for the experiment were recruited and run at local schools, and the head-mounted eye-tracking system was more easily portable than the remote system. This technical difference between Experiments 2 and 3 does not present a problem for comparison of the results. The two systems are produced by the same manufacturer, share the same technical characteristics and can be used interchangeably. The ISCAN head-mounted eye-tracking visor 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 (for details see Trueswell et al., 1999). 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. Three experimenters were engaged in different tasks. A trained undergraduate research assistant presented the experimental pictures in hard copy to the child and interacted with her during the study. A graduate 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 watched the picture image during the experiment, evaluated the eye position and adjusted the angle of the eye-tracking monocle as needed. 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 the child underwent an eye-tracking calibration procedure. The child was asked to briefly hold her gaze on a sequence of five spatial positions on an 8”x 11” page; each position contained a colored picture of a familiar animal. Since the children were eager to cooperate, our success rate of 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 her arm reach. The first experimenter was seated close to and facing the child so that she could flip the page in the binder, monitor the child’s behavior during the experiment and provide encouraging feedback. The second experimenter was stationed at a different table behind the child from which she could see the child’s pointing actions and play an appropriate instruction from the laptop computer. The speakers were located 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.

In contrast to the electronic presentation of the experimental pictures to the adult participants in Experiment 2, the pictures in this experiment were printed in color on paper and assembled in a binder. Each trial (see (11) above) 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 was delayed until the third experimenter confirmed that accurate fixation had been obtained. Upon hearing the experimental sentence (*Now point to the picture where the boy has placed the box behind him*) and the subsequent follow-up instruction (*Now point to the boy with the cap on*), the child proceeded to answer them by pointing to appropriate parts of the picture with her finger. After each trial children were given encouraging feedback.

Only two types of data were analyzed in this experiment: pointing responses in the forced-choice task and eye movements. As in Experiment 1, the response data were assembled into 3 x 2 x 2 x 16 matrices that brought together the referent preference (made by 16 children in each of the two versions of the experiment) for three exemplars of the two prepositions and two genders. Two such matrices were prepared, one for each of the conditions, REFL vs. PRON. The data were analyzed parametrically in

---

15 A written consent form was signed by each child’s parent prior to his or her participation in the experiment.
analyses of variance. 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. The reaction times that resulted from pointing were not collected due to the highly subjective nature of such actions.

Results. Responses. The preferences of sentence-internal and sentence-external referents for the two pronominal types in the response data obtained for children in Experiment 3 are quite different from the preference choices for adults in Experiments 1 and 2 (Table 5).

Table 5. Children Eye-Tracking Experiment 3: Overall sentence-internal referent preference (%)

<table>
<thead>
<tr>
<th>Percentage referent preference for the pronominal</th>
<th>Sentence-internal</th>
<th>Sentence-External</th>
</tr>
</thead>
<tbody>
<tr>
<td>REFL</td>
<td>94</td>
<td>6^a</td>
</tr>
<tr>
<td>PRON</td>
<td>93</td>
<td>7</td>
</tr>
</tbody>
</table>

^a Incorrect responses.

Children’s referent preferences were identical in both REFL and PRON conditions (Fs<1); they overwhelmingly chose the sentence-internal referent (the boy). The seven percent sentence-external preference is nearly the same as the percent of errors in the REFL 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, we can conclude that while the PRON sentences are ambiguous for adults, they are totally unambiguous for children, with the sentence-internal referent as the only possible antecedent for the SDP.

Eye movements. As with the adult eye-tracking data from Experiment 2, the experimental sentences were divided into three regions (see Figure 6 above). Region 3, the region of interest, started at the onset of the pronominal element (at approximately 2475 ms from the beginning of the sentence). In order to make a direct comparison with the adult eye movement data, the stop point was the same as it had been for adults (at 4475 ms). Note that it took the children longer to do the pointing: at 4475 ms, our chosen stop point, 66% of the children had pointed to one of the pictures, while 96% of the adults had pushed the button. We decided not to analyze the right tail of the eye movement distribution because of great variability in the remaining data (34% for children, 6% for adults). For the coarse-grain analysis, we analyzed the proportion of fixations on the picture consistent with the sentence-internal interpretation compared to the picture with the sentence-external interpretation.

Fig. 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 pronouns than with reflexives, 56% vs. 61%.
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), in contrast to Experiment 2, there was no difference in looks to the sentence-external picture in the PRON condition compared to that in the REFL condition (55% vs. 57%, F<1). But crucially, Segment 2 (3475–4475 ms) revealed a significant difference between the PRON and REFL conditions: the children indeed looked substantially less at the sentence-internal picture in the former, 43% vs. 62% (F(1,15)=11.33, p=.0042). 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 pronoun. However, closer towards the end of the trial, during the second 1000 ms, the children start to notice the referential ambiguity in the PRON 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 5 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 the eye movements of the adults and children revealed different real-time course of referential ambiguity resolution. Adults quickly notice the referential ambiguity and start resolving it online, 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 PRON condition. Children, on the other hand, are much slower in noticing the ambiguity; they operate at a chance level about which picture to choose during the initial 1000 ms after the onset of the pronoun. 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 to preference choice.

Fig. 7. Children’s Eye-Tracking Experiment 3: Probability of fixating the Sentence-Internal referent picture in two conditions (%), for the entire Region 3
5. 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 in the linguistic literature (Hestvik, 1992; Tenny, 1999). However, while the grammar permits both sentence-internal and sentence-external reference, and the visual context is consistent with both, adults readily access the former. It is clear that 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, since such factors as accessibility, order-of-mention, recency, and 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 additional 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 and Wexler, 1999), 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 sentence comprehension are costly. Processing of reflexives like himself, 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 language processor has to access information beyond that provided in the syntax, i.e., discourse.

Second, our experiments, as well as the results from on-line processing of reflexives by Runner, Sussman, and Tanenhau (2001), demonstrate that the eye-tracking method is a feasible probe for testing binding theories and is especially suited to accessing speakers’ intuitions under complex discourse conditions. Recording of eye movements allows for better understanding of what people are doing while resolving referential ambiguities on-line. Previous eye-tracking studies (Tanenhau et al., 1995; Trueswell et al., 1999; Altmann and Kamide, 1999) found evidence for rapid and 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 experiments with adults support a constraint-based model of language processing, in which multiple sources of information are used on-line to guide referential processing. Visual context appeared to be a compelling source of information in our experiments. 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, second-mentioned order, or its distance from the pronoun, as it was chosen a robust 20% of the time in both experiments.

How did children’s processing of the referentially ambiguous pronouns differ from that of adults? The use of the eye-tracking methodology allowed us to uncover a major difference between children’s off-line referent preference and their on-line eye movements data. If we had relied only on the referent preference data, we would have had to conclude that children failed to access the sentence-external interpretation, although both sentence-internal and external interpretations are available in the adult grammar. Our eye movement results clearly show that children can assign sentences with referentially ambiguous SDPs an interpretation corresponding to the sentence-external one. Children’s preference for the sentence-internal referent can be readily explained. The adults strongly preferred the sentence-internal referent, and the children used the same default as adults, although in their case this default was even stronger. The difficulty that children experienced with the sentence-external referent is well attested in the acquisition literature. Avrutin (1994) argues that children face difficulties in all areas of anaphora resolution in discourse. Thornton and Wexler (1999) show 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, 1990; Goodluck, Terzi, and Chocano Díaz, 2001). 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 ability.

The notion of the child’s underdeveloped processing mechanism has been proposed to account for many cases of nonadult behavior in establishing pronominal reference. Grodzinsky and Reinhart (1993) proposed that children’s failures in Principle B contexts arise from the child’s processor being unable to consider two representations, the bound variable and the accidental coreference representation, at once. Reinhart (1999) suggests that computing the accidental coreference requires a child to construct two representations, keep both in working memory and then select the appropriate one. This computation exceeds the child’s processing ability, since her working memory is not yet as developed as the adult’s. Trueswell et al. (1999) and Snedeker, Thorpe, and Trueswell (2001) show that difficulty with utilizing referential information is not due to a general lack of knowledge that the discourse reference exists; instead, it is due 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 the contextual facts further drive processing decisions. And, indeed, our eye movement data demonstrate the children’s awareness of the referential ambiguity. Contrary to the above-mentioned accounts, our results provide evidence that children can hold multiple referential representations in on-line processing and given time, access the sentence-external referent. Moreover, children appeared not to be limited in their ability to process referentially ambiguous pronouns since they clearly entertained the uncommon alternative.

Thus, our results contribute to an important new line of research that strives to establish similarities and differences between the adult and child language processing system. The null hypothesis is that the underlying architecture of the child’s language processing system is essentially the same as the adult system (Goodluck and Tavakolian, 1982; Fodor, 1998; Thornton and Wexler, 1999), a processing extension of the Continuity hypothesis in acquisition. The growing body of literature about on-line processing in children is putting this hypothesis to the empirical test. Some of the existing studies are compatible with it (Nadig and Sedivy, 2000; Booth, MacWhinney, and Harasaki, 2000; Traxler, 2001), while others find important differences: children’s inability to utilize the contextual cues and to revise the initial incorrect interpretation (Trueswell et al., 1999; Snedeker et al., 2001), their non-adult-like processing of accessibility (Arnold et al., 2000), their insensitivity to prosodic cues in the relative clause attachment ambiguities (Smyth, 2001) and their different second-pass parsing (Hahne and Friederici, 2001).

We have demonstrated that children’s eye movement data were qualitatively similar to adults’, showing their emerging adult-like awareness of the referential ambiguity in sentences with SDPs. 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. Instead, it is the fact that children need more time for accessing discourse reference and integrating discourse information during sentence processing than adults do. Children 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. In addition, the discrepancy between the referent preference and eye movement data may be taken as an argument that although children may have the same parser as adults, their parser does not work as efficiently as it does for adults. Children did not revise their initial commitment to the sentence-internal referent, thus displaying deterministic rather than probabilistic parsing. This is similar to the Trueswell et al. (1999) results in which children were unable to revise their initial commitment to an interpretation that turned out to be incorrect when, later in the sentence, they encountered incompatible linguistic information. Note, however, that in our experiment there was no incompatible information later in the sentence: the sentences were globally ambiguous, and both referents were potentially possible. The lack of urgent necessity to backtrack allowed the children to rely on the ‘minimal commitment’ strategy and hold on to the sentence-internal interpretation. As the human processing system matures, both obligatory optional revisions of a minimal commitment parsing become increasingly faster. Determining the precise developmental stage at which children begin to parse probabilistically remains a topic for future research.
REFERENCES


APPENDIX

Spoken materials for the experiments

Differences in the materials for the three experiments.

1. The set of materials to follow represents a full Version 1 of Experiment 2, “Adult Eye-Tracking”.
2. Experiment 1, “Referent Preference Questionnaire”, consisted only of 24 experimental items and no fillers. Each trial had only two sentences in it: the preamble sentence and the experimental question.
3. Experiment 3, “Children Eye-Tracking”, represented a proper subset of 32 trials from Experiment 2, “Adult Eye-Tracking”, and contained 16 experimental trials (6, 11, 18, 19, 29, 32, 33, 36, 37, 39, 42, 43, 45, 47, 50, 52) and 16 fillers (1F, 2F, 3F, 5F, 9F, 12F, 15F, 20F, 22F, 28F, 30F, 34F, 38F, 40F, 44F, 51F, 54F). In addition, children’s trials were shorter than the adults’, consisting of only four sentences per trial.

Explanations of the materials.

1. The letter F after the number of the trial designates a filler.
2. The letters [R] and [L] designate the right or the left panel as the correct answer in the filler and the experimental trials with the reflexive, and the expected preferred interpretation in the experimental trials with the pronoun.
3. Sentences 2, 4 and 5 in each trial were used as additional fillers in the eye-tracking Experiments 2 and 3.
4. The trials whose numbers are printed in boldface are the experimental trials. They were rotated between the two experimental conditions, REFL or PRON, depending on the version of the experiment.

1F [R].
In these pictures, there are some children.
[Now look at the cross.]
Which picture shows that the girl is drying herself?
Now look at the girl's turquoise towel.

2F [L].
In these pictures, you can see a dog, a painting, and a woman.
[Now look at the cross.]
Which picture shows that the dog was pointing to him?
Now look at the dog's tail.

3F [L].
In these pictures, you can see a woman and a boy.
[Now look at the cross.]
Which picture shows that the woman is waving?
Look at the woman's green skirt.
Now look at the boy's balloon.

4F [R].
In these pictures, you see two boys and some animals.
[Now look at the cross.]
Which picture shows that the horse and the pig are running?
Look at the pig's tail.
Now look at the boy with the hat on.

5F [R].
In these pictures, there are some boys on elephants.
[Now look at the cross.]
Which picture shows that two boys were riding on elephants?
Now look at the boy with the yellow shirt.
6 [him-R].
In these pictures, you see a boy, a grandfather, and a bucket.
[Now look at the cross.]
Which picture shows that the grandfather has put the bucket behind him?
Now look at the red bucket.

7 F [L].
In these pictures, you can see elephants and balloons.
[Now look at the cross.]
Which picture shows that there are three elephants holding balloons?
Look at the orange balloon.

8 [him-R].
In these pictures, you see a boy, a man, and a hairbrush.
[Now look at the cross.]
Which picture shows that the father has placed the hairbrush near him?
Now look at the boy with the orange shirt.

9 F [R].
In these pictures, you can see some girls, a woman, a man, and a painting.
[Now look at the cross.]
Which picture shows that the girls are pointing at her?
Now look at the woman.

10 F [R].
In these pictures, you can see a dinosaur and a boy at a desk.
[Now look at the cross.]
Which picture shows that the dinosaur carried a flag?
Look at the dinosaur's red and white flag.
Now look at the boy's cap.

11 [herself-L].
In these pictures, there is a woman, a girl, and a box.
[Now look at the cross.]
Which picture shows that the girl has placed the box near herself?
Now look at the red box.

12 F [L].
In these pictures, you can see some horses.
[Now look at the cross.]
Which picture shows that the boy is riding a brown horse?
Now look at the child on the other horse.

13 [himself-R].
In these pictures, there is a boy, a man, and a cup.
[Now look at the cross.]
Which picture shows that the boy has put the cup behind himself?
Now look at the boy with the green shirt.

14 F [L].
In these pictures, there are some giraffes.
[Now look at the cross.]
Which picture shows that the giraffe is standing alone?
Look at his tail.
Now look at the red balloon.
15F [R].
In these pictures, you can see some boys and some umbrellas.
[Now look at the cross.]
Which picture shows that one boy is holding an umbrella?
Now look at that boy's shirt.

16 [her-R].
In these pictures, you see a girl, a woman, and a horse.
[Now look at the cross.]
Which picture shows that the girl has placed the horse near her?
Now look at the girl wearing a red shirt.

17F [L].
In these pictures, you can see children on horses.
[Now look at the cross.]
Which picture shows that two children were riding horses?
Look at the child with the cap on.

18 [herself-L].
In these pictures, there is a girl, a mother, and a hairbrush.
[Now look at the cross.]
Which picture shows that the mother has put the hairbrush behind herself?
Now look at the woman with the pink dress.

19 [him-L].
In these pictures, there is a man, a boy, and a suitcase.
[Now look at the cross.]
Which picture shows that the boy has put the suitcase near him?
Now look at the boy with the blue cap.

20F [R].
In these pictures, there are some elephants holding balloons.
[Now look at the cross.]
Which picture shows that the elephants are holding red balloons?
Look at one of the green balloons.

21 [herself-R].
In these pictures, you can see a woman, a girl, and a chair.
[Now look at the cross.]
Which picture shows that the girl has placed the chair behind herself?
Now look at the brown chair.

22F [L].
In these pictures, you see a man, a woman, a boy, and a girl.
[Now look at the cross.]
Which picture shows that the man is in the tub?
Look at the bath sponge.
Now look at the woman.

23F [L].
In these pictures, there are two girls, two boys, and two pigs.
[Now look at the cross.]
Which picture shows that the girls are wrapped in towels?
Look at the girl with black hair.
Now look at the tallest boy.
24 [herself-R].
In these pictures, you see a girl, a woman, and a watering can.
[Now look at the cross.]
Which picture shows that the girl has put the watering can near herself?
Now look at the green watering can.

25F [R].
In these pictures, there are some boys and umbrellas.
[Now look at the cross.]
Which picture shows that the boy was holding three umbrellas?
Look at the red umbrella.
Now look at the boy with the striped shirt.

26F [L].
In these pictures, you can see a man, a child, and a mirror.
[Now look at the cross.]
Which picture shows that the man is looking at himself?
Look at the man's hat.
Now look at the child's hair.

27 [him-L].
In these pictures, you can see a boy, a man, and a box.
[Now look at the cross.]
Which picture shows that the boy has placed the box behind him?
Now look at the boy with the cap on.

28F [R].
In these pictures, you can see a man and a giraffe.
[Now look at the cross.]
Which picture shows that the giraffe was looking at the painting?
Look at the dog in the painting.
Now look at the man's tie.

29 [her-R].
In these pictures, there is a grandmother, a girl, and a cup.
[Now look at the cross.]
Which picture shows that the girl has put the cup near her?
Now look at the woman with the blue skirt.

30F [R].
In these pictures, you can see two women and two girls.
[Now look at the cross.]
Which picture shows that the girl is washing herself?
Look at the bubbles.
Now look at the girl's skirt.

31F [L].
In these pictures, there are two pigs.
[Now look at the cross.]
Which picture shows that the pig has a red apple in its mouth?
Look at that pig's tail.
32 [himself-L].
In these pictures, you can see a man, a boy, and a horse.
[Now look at the cross.]
Which picture shows that the boy has placed the horse behind himself?
Now look at the man with the beard.

33 [him-R].
In these pictures, there is a man, a boy, and a watering can.
[Now look at the cross.]
Which picture shows that the boy has put the watering can behind him?
Now look at the green watering can.

34F [L].
In these pictures, there are dinosaurs and balloons.
[Now look at the cross.]
Which picture shows that there are two dinosaurs holding balloons?
Look at the dinosaur with the green balloon.

35F [L].
In these pictures, you can see giraffes and balloons.
[Now look at the cross.]
Which picture shows that two giraffes are holding balloons?
Now look at the giraffe that is standing alone.

36 [herself-R].
In these pictures, you can see a girl, a grandmother, and a bucket.
[Now look at the cross.]
Which picture shows that the grandmother has placed the bucket near herself?
Now look at the red bucket.

37 [himself-L].
In these pictures, you can see a boy, a father, and a bear.
[Now look at the cross.]
Which picture shows that the father has put the bear near himself?
Now look at the boy without a hat.

38F [L].
In these pictures, there is a man, some boys, and a horse.
[Now look at the cross.]
Which picture shows that the boy is scratching himself?
Look at the man.
Now look at the horse in the painting.

39 [her-L].
In these pictures, you can see a girl, a woman, and a ball.
[Now look at the cross.]
Which picture shows that the girl has placed the ball behind her?
Now look at the red ball.

40F [R].
In these pictures, you can see some children and some adults.
[Now look at the cross.]
Which picture shows that the girl was waving to him?
Now look at the woman who is turned away from us.
In these pictures, you can see giraffes and balloons. Which picture shows that the giraffe was holding just one balloon? Now look at the giraffe that is holding three balloons.

42 [himself-R]. In these pictures, you can see a boy, a grandfather, and a ball. Which picture shows that the boy has placed the ball near himself? Now look at the blue ball.

43 [herself-R]. In these pictures, there is a girl, a grandmother, and a suitcase. Which picture shows that the girl has put the suitcase behind herself? Now look at the girl with the blue dress.

44F [L]. In these pictures, you can see a man, a girl, and a boy. Which picture shows that the man looked at her? Now look at the man's hat. Look at the boy's apple.

45 [him-L]. In these pictures, you can see a man, a boy, and a chair. Which picture shows that the boy has placed the chair near him? Now look at the blue chair.

46F [R]. In these pictures, you can see elephants. Which picture shows that the elephants are looking at each other? Now look at the smallest boy.

47 [her-L]. In these pictures, there is a woman, a girl, and a hat. Which picture shows that the mother has put the hat near her? Now look at the girl with the pink dress on.

48 [himself-L]. In these pictures, there is a boy, a grandfather, and a hat. Which picture shows that the grandfather has placed the hat behind himself? Now look at the man with the brown sweater.

49F [L]. In these pictures, you can see three girls and a woman. Which picture shows that both girls have touched their heads? Look at the girl with the black hair.
50 [her-R].
In these pictures, you can see a grandmother, a girl, and a bear.
[Now look at the cross.]
Which picture shows that the grandmother has placed the bear behind her?
Now look at the woman with the pink apron.

51 F [R].
In these pictures, you see many giraffes.
[Now look at the cross.]
Which picture shows that two of them are looking at each other?
Now look at the picture with three giraffes.

52 [himself-L].
In these pictures, you can see a man, a boy, and a pig.
[Now look at the cross.]
Which picture shows that the boy has put the pig near himself?
Now look at the man with the hat on.

53 [her-L].
In these pictures, you can see a woman, a girl, and a pig.
[Now look at the cross.]
Which picture shows that the girl has put the pig behind her?
Now look at the girl with the blue dress.

54 F [R].
In these pictures, there are two women and a girl.
[Now look at the cross.]
Which picture shows that the girl has touched the woman's hair?
Look at the woman’s glasses.
Now look at the broom.