Symmetry in comprehension and production of pronouns:
A comparison of German and Hebrew

Esther Ruigendijk¹, Naama Friedmann², Rama Novogrodsky², and Noga Balaban²
¹Oldenburg University, Department of Dutch, Oldenburg, Germany
²Tel Aviv University, Israel

Abstract

In the course of language acquisition, children interpret pronouns differently from adults. In some languages and ages, children accept “The girl and grandma met, and then the girl combed her” as describing the girl combing herself. This study tested whether this difficulty is symmetrically reflected in comprehension and production, by assessing the comprehension and production of pronouns and reflexives in children acquiring German and Hebrew. Comprehension was tested using a sentence-picture matching task in 44 German-speaking children aged 3;3-6;2, and 54 Hebrew-speaking children aged 2;4-6;7. Production was assessed using a sentence elicitation task for 44 German-speaking children aged 3;1-6;8, and 60 Hebrew-speaking children aged 2;4-6;7. The results indicate a very clear symmetry between comprehension and production of pronouns in language acquisition. German-speaking children master the comprehension of pronouns and do not allow for local antecedents for pronouns already from age 4;3, and they also do not use pronouns for local antecedents in production. Crucially, the Hebrew-speaking children, who allow local antecedents in pronoun interpretation up to age 6, also produce pronouns with local antecedents (namely, to describe reflexive actions) up to this age. An individual-level analysis of the performance of 50 Hebrew-speaking children who were tested in both comprehension and production yielded a similar symmetrical pattern: the children who incorrectly produced pronouns in the reflexive condition showed the same pattern in comprehension, choosing a reflexive picture (with local antecedent) for sentences with pronouns.

Keywords: pronouns, reflexives, acquisition, production, comprehension, German, Hebrew

1. Introduction

1.1. The acquisition of pronouns and reflexives

In the course of language acquisition, children understand pronouns differently from adults (see e.g., Chien and Wexler, 1990 and Wexler and Chien, 1985 on English; Koster, 1993 and Philip and Coopmans, 1996 on Dutch). The main difference between the way children and adults interpret pronouns is that in certain languages children allow pronouns to refer to a local antecedent. Namely, adults who hear the sentence “The fairy and the girl met, and then the fairy pinched her” understand that the fairy is pinching the girl, and would not accept an interpretation according to which the fairy is pinching herself. Conversely, kindergarten children allow both interpretations, sometimes taking the local noun phrase fairy to be the antecedent of the pronoun her. At the same time, children as young as 4 years understand reflexive elements in the same way

⁰ We thank Melanie Kläne and Hendrikje Ziemann for their help in the German tests. This research was supported by the Lieselotte Adler Laboratory for Research in Child Development, and the writing of this manuscript was supported by the COST A33 European initiative (PI Sauerland).
as adults, that is, when hearing a sentence like “The fairy and the girl met, and then the fairy pinched herself”, they usually interpret the reflexive element correctly, that is, referring to the local subject. Some researchers termed this pattern “the Delay of Principle B Effect” (also DPBE), referring to Principle B of Chomsky’s Binding Principles (see (1), Chomsky, 1981). This label, DPBE, suggests that children acquire principle B later than principle A.

(1) Binding Principles (GB, Chomsky, 1981)¹

A: An anaphor is bound within its governing category.

B: A pronoun is free in its governing category.

Stating the Binding Principles in a somewhat simplified way, these principles determine that a reflexive element such as ‘herself’ must find its antecedent in the local domain, whereas a pronoun cannot find its antecedent in the local domain (see Reinhart and Reuland, 1993 for an extensive discussion of these principles and related issues as well as for a further development of Binding Theory into the Reflexivity model).

Recent work has demonstrated that the non-adult pattern of pronoun interpretation varies across sentence constructions and languages, and it was also found to be influenced by the task. Chien and Wexler (1990) showed, for example, that in sentences with a quantified subject, like ‘Every fairy pinched her’, even 4-year-old children reject the local interpretation for the object pronoun ‘her’ most of the times, which suggests that application of Principle B (i.e., a restriction on variable binding) is not in itself problematic. Another type of variation across structures has been found in Dutch, where more problems arise with the interpretation of pronouns in so-called ECM constructions like ‘The girl sees her dancing’ than in transitive structures like ‘The girl pinched her’ (e.g., Baauw, 2002; Philip and Coopmans, 1996).²

Another source of variation relates to languages with different characteristics. In some languages, most notably Romance languages, children do not show the described interpretation problems with pronouns in the transitive construction (see for French: Hamann, 2002; Hamann et al., 1997; Jakubowicz, 1984; for Italian: McKee, 1992; Catalan: Escobar and Gavarró, 1999; and for Spanish: Baauw, 2002). Problems with the interpretation of pronouns in ECM constructions were reported in these languages too (French: Hamann, 2002; Hamann et al., 1997; Catalan: Escobar

¹ Where α is bound by β if β c-commands α and is co-indexed with it and y is a governing category for α if and only if y contains α, a governor for α, and an accessible subject. Governing category can be loosely defined as local domain, for a precise definition see Chomsky (1986).

² There are several accounts for the fact that in ECM constructions pronoun interpretation problems are even stronger than in simple transitive constructions. Since a discussion of these accounts is beyond the scope of this paper, the reader is referred to Baauw (2002), Philip and Coopmans (1996), Reinhart and Reuland (1993), Reuland (2001), or Vasić (2006).
The problems found in comprehension are not said to occur in production (see e.g., Bloom et al., 1994; De Villiers, Cahillane, and Altreuter, 2006; Spenader, Smits, and Hendriks, 2009). Bloom and others (1994) analyzed the production of first person pronouns and reflexives in the CHILDES spontaneous speech database (McWhinney, 2000), and found that children realized both elements correctly most of the times. De Villiers and colleagues (2006) tested the production of third person pronouns and reflexives by English-speaking children (with a mean age of 6;3) using a picture description task. They reported almost perfect production of pronouns. The children in their study showed the classical pronoun comprehension problems, but never produced reflexives when a pronoun was required, but they did use full DPs and proper names instead of pronouns, which were grammatical alternatives. Unfortunately, results on the reflexive condition (i.e., a picture that depicts a reflexive action), which are the crucial condition for testing for symmetry in the choice of antecedents for pronouns, were not reported. Spenader and others (2009) did a similar experiment with Dutch children and claimed to find good performance on the production of both pronouns and reflexives. However, some indications for incorrect use of pronouns can be found in their data. We will return to this in the discussion.

To summarize, whereas children speaking languages like Dutch and English may still interpret ‘her’ as co-refering with ‘the girl’ in ‘The girl pinched her’, these mistakes, overusing pronouns for a reflexive meaning, have not been reported for production. However, while comprehension has been tested thoroughly in a structured way, across languages and sentence types, data regarding the classical ‘delay of principle B effect’ in production of pronouns and reflexives, are scarce, at least to our knowledge. The question of this study is therefore: is there really an asymmetry between comprehension and production of pronouns?

Data from spontaneous speech are reported (e.g., Bloom et al., 1994), but if children have difficulties with pronouns, they may resort to avoidance, instead of using the incorrect form. The same applies, to a lesser extent, for picture description. We therefore believe that structured elicitation tests are more appropriate for testing children’s ability to realize pronouns in production.

---

3 The better performance of children in Romance languages has often been attributed to the fact that these languages have pronouns that are syntactic clitics, which can occur VP-externally (Baauw, 2002; McKee, 1992) and only exceptionally allow accidental coreference or guise creation (see Avrutin, 1994; Hamann et al., 1997; and Thornton and Wexler, 1999 for details).

4 Costa et al. (2008) used a structured elicitation task to test pronoun and reflexive production in European Portuguese. Their research question was, however, whether children produce pronouns and reflexives or whether they use the null object option, and hence omit them, rather than a question about the antecedent used.
We suggest that in order to test whether children assign a local antecedent to pronouns in production, as they do in comprehension in some languages, the target sentences that should be elicited are not the ones that require pronouns in adult speech. In fact, if we want to see whether children assign local antecedents for pronouns in production, reflexive actions should be tested. If a child describes a situation in which an elephant washes himself as ‘The elephant washes him’, we will have the parallel for the pattern found in comprehension.

A further way to test for symmetry between comprehension and production is to compare the pattern of production in languages that show different patterns of comprehension with respect to the acquisition of pronoun interpretation, and see whether the production shows similar patterns: whether the language in which there is no pronoun interpretation problem in comprehension also shows no problem in production, and in the language in which children allow a local antecedent in comprehension, children also use pronouns to refer locally in production. For this purpose, we compared in this study data on comprehension and production from two languages that show different patterns of pronoun comprehension in acquisition: German and Hebrew.

1.2. German vs. Hebrew

The reason we compared comprehension and production in German and Hebrew is that differences in the pronoun comprehension pattern have been reported for these two languages. Children who acquire German do not show the classical pattern of pronoun interpretation (Ruigendijk, 2008) and hardly ever assign a local antecedent to the object pronoun at the age of 4. In contrast, Hebrew speaking children assign pronouns a local antecedent until they are 6 years old (Friedmann, Novogrodsky, and Balaban, in press). Therefore, it will be interesting to compare the performance on elicited production of pronouns in these two languages to find out whether there is symmetry between comprehension and production.

Hebrew and German are furthermore suitable languages for exploring this question, since very similar surface structures can be tested, as can be seen in examples (2) for reflexives and (3) for pronouns. In both languages, third person pronouns are inflected for gender and number. In Hebrew, third person reflexives are also inflected for gender and number.

(2) Hebrew: ha-yeled, meshabeax et acmoi
   German: der Junge, lobt sich
   the boy, praises SE

(3) Hebrew: ha-yeled, meshabeax oto*i/j
   German: der Junge, lobt ihn*i/j
   the boy, praises him*i/j

See Ruigendijk (2008) for possible explanations of the comprehension pattern in German.
2. Methods

2.1. Comprehension task

2.1.1. Material

Comprehension was tested with a picture selection task. The German task included 16 items per participant, half with a reflexive and half with a pronoun. The Hebrew task included 24 sentences per participant, half with a reflexive and half with a pronoun. The sentences with pronouns and reflexives were presented in a semi-randomized order. Examples of sentences used in the picture selection task in German and Hebrew are given in (4) and (5) respectively.

(4) Erst haben die Frau und die Kleine gelaufen, und dann hat die Frau sich/sie gewaschen
‘First the woman and the girl walked and then has the woman SE/her-ACC washed’

(5) Ha-yalda ve-ha-jirafa nifgeshu, ve-az ha-yalda raxca et acma/ota
‘The girl and the giraffe met and then the girl washed ACC self-her/ACC-her’

For both languages, half of the sentences contained two feminine DPs and half contained two masculine DPs. The two DPs in each sentence always had the same gender in order to preclude gender agreement cue for choice of antecedent. In the German task, the two antecedents (i.e. ‘die Frau’ and ‘die Kleine’ in (4) were introduced in the first conjunct which was accompanied with a picture. For the second conjunct, three pictures were presented, as shown in Fig. 1, out of which the child had to choose the one that matched the sentence. The three pictures included the matching picture, a picture in which the object was the incorrect antecedent, and a lexical distractor with the same DPs but a different verb (which was included originally to allow for direct comparison of this study to studies with aphasics. This lexical foil was not selected by any of the children). The Hebrew sentences were presented with one page that included two pictures, one matched the sentence, the other included the incorrect antecedent as the object, as seen in Fig. 2 (the introduction of the possible antecedents in the first conjunct and the second conjunct was uttered when these two pictures were in front of the child, without an additional introductory picture).

---

6 The German comprehension task included 6 additional conditions: pronouns and reflexives in ECM and relative constructions, and subject and object pronouns in conjoined sentences. The Hebrew comprehension task also included 4 additional conditions: reflexives and pronouns in sentential complements and in subject relative constructions. These conditions were not relevant to the current comparison.
2.1.2. Procedure

The sentences were orally presented by a native speaker of the relevant language in a neutral intonation and at a normal rate. The sentences were presented without time limitation, and were repeated as many times as the child requested or needed. If the child changed her response, only the final answer was included in the calculation. The children were tested individually in a quiet room.
2.2. Production task

Production was examined in both languages with the same elicitation task. The test included 16 items per participant, the target object of 8 of them was a pronoun, and for 8 it was a reflexive, as seen in (6) and (7). Sentences were presented in semi-randomized order, with the same order for both languages.

(6) pronoun target:
Here the woman wets the girl, and what does the woman do to the girl here?
\(\Rightarrow\) (she) dries her.

(7) reflexive target:
Here the boy combs the horse, and what does the boy do here?
\(\Rightarrow\) (he) combs himself.

In the production task too, the two figures in each item had the same gender. All the verbs in the production task were transitive. Half of the verbs in the reflexive condition also had a reflexive form in Hebrew that is created by a morpho-phonological operation. However, the task was constructed in such a way that the contrast between the lead-in picture and sentence and the target picture would encourage the use of a reflexive pronoun rather than the reflexive verb.

Each item was presented with two pictures, one above the other, as shown in Fig. 3, which accompanied the sentence in example (6), and Fig. 4 which accompanied the sentence in example (7). We elicited the sentences by providing the lead-in sentences, as exemplified in (6) and (7), pointing, first to the upper picture, and then, with the question, to the picture on the bottom, and the child answered the question.

Fig. 3. A picture that was used in the pronoun elicitation condition.
The whole session was audio-recorded and transcribed during and after the session. Each response was scored as either target or non-target response. The non-target responses were categorized in the following way: incorrect pronominal element (pronoun instead of a reflexive or vice versa); full DP instead of the pronominal element; omission of the pronominal element; no response. Responses in which a different verb was used that preserved the argument structure of the target verb (such as using “washing” or “petting” instead of “soaping”) were calculated together with responses that included the target verb. The use of an intransitive verb or a completely irrelevant response was scored as “other description”.

2.3. Participants

For the German tasks, the comprehension task was carried out with 44 children (range 3;3-6;2; 27 of them were girls; 6 children were tested with only half of the items). The group was divided into four age groups: 3-year olds (n=9, mean age 3;6, SD 0;2), 4-year olds (n=17, mean age 4;6, SD 0;3), 5-year olds (n=14, mean age 5;6, SD 0;3), and 6-year olds (n=4; mean age 6;2, SD 0;1). The production task was presented to another group of 44 children (range 3;1-6;8, 20 of them were girls, 6 children were tested both with the comprehension and the production task). This group too was divided into four age groups: 3-year olds (n=12, mean age 3;6, SD 0;3), 4-year olds (n=11, mean age 4;6, SD 0;3), 5-year olds (n=11, mean age 5;6, SD 0;3), and 6-year olds (n=10; mean age 6;4, SD 0;4).
For Hebrew, the *comprehension* task included 54 children aged 2;4-6;7, 28 of them were girls. This group was divided into 5 age groups: 2-year olds (n=7, mean age 2;8, SD 0;3), 3-year olds (n=12, mean age 3;5, SD 0;3), 4-year olds (n=11, mean age 4;7, SD 0;3), 5-year olds (n=14, mean age 5;5, SD 0;3), and 6-year olds (n=10; mean age 6;4, SD 0;2). The *production* task included 60 children aged 2;4-6;7, 29 of them were girls. This group was divided into 5 age groups: 2-year olds (n=7, mean age 2;8, SD 0;3), 3-year olds (n=14, mean age 3;5, SD 0;3), 4-year olds (n=15, mean age 4;5, SD 0;3), 5-year olds (n=14, mean age 5;6, SD 0;3), and 6-year olds (n=10; mean age 6;4, SD 0;2). To allow for direct comparison between comprehension and production on the individual level, 50 of the Hebrew-speaking children participated in both tasks.

3. Results

The main findings are that German children do not have problems in the comprehension of either pronouns or reflexives after age 4, and that, symmetrically, they also do not produce pronouns in the reflexive condition. Hebrew-speaking children, on the other hand, clearly show non-adult-like pronoun interpretation until the age of 6. Importantly, when they allow a local antecedent for the pronoun in comprehension, they also make errors on the reflexive items in the elicitation task, describing them with a pronoun.

3.1. German Results

The German results are presented in Tables 1 and 2 and in Figures 5 and 6. Fig. 5 shows that German children do not exhibit the pattern of pronoun interpretation problems that has been reported for example for English and Dutch. Four-year old German children already do not have problems in the comprehension of either pronouns or reflexives. In fact, from age 4;3, all the German-speaking children performed either flawlessly or with a single error on both the reflexive and the pronoun conditions, namely, all of them were above chance. Three-year olds do make some errors on the comprehension task, but both with reflexives and with pronouns, hence they also do not exhibit the typical pattern. Even the three-year-olds did not make many errors: 8 of them made 2 or less errors in the pronoun condition, and 7 of them made 2 or less errors in the reflexive condition.
Because we see a natural cut-off point in performance at age 4;3, Table 1 presents percentages correct and the number of children who performed above chance (7 or 8 out of the 8 items) in the different age groups respecting this cut-off.

Table 1. German: number of participants who performed above chance in the comprehension of reflexives and pronouns, and the groups’ percentage correct

| Participants above chance / number of participants | % Correct | | | |
|---|---|---|---|---|---|
|  | Reflexive | Pr. | Reflexive | Pr. |
| 3;3-4;2 | 4/12 | 6/12 | 74% | 79% |
| 4;3-6;3 | 32/32 | 32/32 | 97% | 97% |

The German-speaking children showed similar performance on reflexives and pronouns at each of the ages tested: \( t(8) = 0.63, p = .54, t(16) = 1.38, p = .19, t(13) = 0, p = 1, t(3) = 2.32, p = 10, \) for the 3, 4, 5, and 6 year olds respectively. Their performance improved from age 3 to age 4 on both reflexives and pronouns (marginally significant for the reflexives, \( U = 47.5, p = .06, \) and significantly for the pronouns \( U = 30.5, p = .007 \)). After age 4 (or more specifically for the group tested, age 4;3), their performance on both reflexives and pronouns no longer changed. They showed no

---

7 Because most of the 3-, 4- and 5-year old participants had exactly the same performance on pronouns and reflexives, the analysis included too many ties, and a Wilcoxon analysis could not reach significance anyway, and therefore a \( t \)-test was used. A \( t \)-test was also used for the comparison of the 5- and the 6-year-old participants, because of the small number of 6-year old participants, which did not allow the use of the Mann-Whitney test.
significant difference between the performance of the 4- and the 5-year olds on both reflexives \((U = 98, p = .21)\) and pronouns \((U = 106.5, p = .32)\), and no difference between the performance of the 5- and the 6-year olds on either reflexives or pronouns \((t(16) = 0.36, p = .36\) for both comparisons).

Symmetrically, as can be seen in Table 2, they also hardly ever produced pronouns in the reflexive condition. Recall, that this is the crucial condition for testing the symmetry question: they did not assign a local antecedent to pronouns in comprehension, nor in production.

As shown in Table 2, not all the responses to the elicitation task were target-like. Crucially, however, the non-target responses hardly ever included using the reflexive instead of the pronoun. The non-target responses did include answers that could be ascribed to the nature of the task, such as the “other descriptions”, which occurred mainly in the youngest groups, and to difficulties related to discourse demands, which reflected in object omissions and the use of full DPs instead of pronouns.

### 3.2. Hebrew results

The acquisition of Hebrew pronouns shows a pattern that is completely different from German, in both comprehension and production. In comprehension, as shown in Fig. 6, whereas Hebrew-speaking children understood reflexives from age 4, they only started to interpret pronouns correctly when they were six years old. With respect to reflexives, as shown in Table 3, from age 4 all the Hebrew-speaking children perform above chance on the reflexive condition, without a significant difference between the 4- and the 5-year olds, or between the 5- and the 6-year olds \((U = 87.5, p = .29; U = 71.5, p = .48,\) respectively). The finding that from age 4 they started understanding reflexives was also reflected in the fact that a significant difference was found between the 3- and the 4-year-olds, \(U = 102, p = .01\).

The pronoun comprehension in Hebrew showed a very different pattern of comprehension from the reflexives. Not even all six year olds performed above chance level on pronoun interpretation, as shown in Table 3. There was no significant difference between the 3- and the 4-year olds with respect to pronoun comprehension, and no difference even between the 4- and the 5-year olds \((U = 70.5, p = .40; U = 74, p = .44,\) respectively). The 6-year olds understood pronouns significantly better than the 5-year olds, \(U = 103.5, p = .03\).
Fig. 6. Hebrew Comprehension: Percentage correct on picture selection

Table 3. Hebrew: number of participants who performed above chance in the comprehension of reflexives and pronouns, and percentage correct in each age group

<table>
<thead>
<tr>
<th>Participants above chance / number of participants</th>
<th>Reflexive</th>
<th>Pronoun</th>
<th>% Correct Reflexive</th>
<th>% Correct Pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1/7</td>
<td>1/7</td>
<td>65%</td>
<td>65%</td>
</tr>
<tr>
<td>3</td>
<td>7/12</td>
<td>4/12</td>
<td>74%</td>
<td>66%</td>
</tr>
<tr>
<td>4</td>
<td>11/11</td>
<td>3/11</td>
<td>93%</td>
<td>68%</td>
</tr>
<tr>
<td>5</td>
<td>14/14</td>
<td>4/14</td>
<td>95%</td>
<td>65%</td>
</tr>
<tr>
<td>6</td>
<td>10/10</td>
<td>7/10</td>
<td>95%</td>
<td>83%</td>
</tr>
</tbody>
</table>

This pattern, of reflexives being acquired around the age of four, and pronouns being acquired around age six, is also reflected in pairwise comparisons between reflexives and pronouns: No difference exists yet between reflexives and pronouns at age 3, an age in which Hebrew-speaking children still perform poorly on both, \( T = 22, p = .18 \). Then, at ages 4 and 5, the comprehension of reflexives was significantly better than the comprehension of pronouns (\( T = 4, p = .003; T = 0, p < .001 \), for ages 4 and 5 respectively). At age six this difference diminishes, because of the relatively improved performance on pronouns, and becomes only marginally significant, \( T = 5, p = .08 \).

The tendency of the Hebrew-speaking children to assign a local antecedent to pronouns in comprehension is symmetrically manifested in their responses in the elicited production task. As shown in Table 4, Hebrew-speaking children of all ages, even the six year olds, produced pronouns
in the reflexive condition. Examples for the use of a pronoun instead of a reflexive are provided in (8) and (9).

(8) Experimenter (presenting the picture-pair shown in Fig. 2, in which the bottom picture depicts a girl soaping herself): Here the girl is soaping a giraffe. What does she do here?
   Ro. 3;4: soaping her.

(9) Experimenter (showing a picture pair in which on top, a child is cleaning a teddy-bear, and below, the child is cleaning himself): Here the boy is cleaning a teddy-bear. What does he do here?
   Do. 6;3: cleaning him.

Apart from pronouns, some other non-target responses were given, including some “don’t know” responses and object omissions, which occurred mainly in the youngest groups. The distribution of all types of responses to the reflexive condition is shown in Table 4.

In the pronoun condition, reflexives were hardly ever used instead of pronouns. As in German, some other non-target responses were produced, which, as shown in Table 4, included mainly other descriptions and no responses, especially in the younger age groups, as some full DPs and omissions, which occurred less frequently than in German. Crucially, none of these responses included a wrong choice of antecedent. It should be mentioned that both in Hebrew and in German, full DPs were non-target for the task, but nevertheless grammatical.

3.3. Symmetry between comprehension and production at the single participant level

Another informative way to assess the symmetry between comprehension and production is to test performance of individual participants and see whether the comprehension and production problems co-occur, namely, whether those participants who still assign the local antecedent to a pronoun in comprehension are also those who use pronouns in the reflexive condition in production. For this aim, we tested 50 of the Hebrew-speaking participants in both comprehension and production. The results were very clear-cut: of the 50 children, 24 produced at least one pronoun in the reflexive condition. All of these children, but one 2;8 girl, did not master comprehension of pronouns yet, as reflected in their choice of reflexive pictures for sentences with pronouns on two or more items. Testing the other direction resulted in a similar convincing pattern, namely, 13 of the children showed mastery of pronoun interpretation, performing either flawlessly or with a single error in the comprehension task (out of the 12 sentences). None of
these children used pronouns with a local antecedent in the production task (with the exception of the same 2;8 year old girl8).

3.4. Summary – comparison of German and Hebrew in comprehension and production

Fig. 7 provides a direct comparison between German and Hebrew with respect to the comprehension of pronouns and reflexives in ages 4 and 5, the ages that demonstrate the asymmetry between the languages and between the pronominal elements most clearly; Fig. 8 describes the production of pronouns in the reflexive condition in the two languages.

The difference shown in these figures with respect to the comprehension and production of pronouns is very clear: German-speaking children master the comprehension of pronouns already at age 4, children who speak Hebrew allow local antecedents in pronoun interpretation up to age 6. Crucially, this same pattern is symmetrically reflected in their production: German children do not produce pronouns with a local antecedent, Hebrew-speaking children do.

Fig. 7. Comprehension of reflexives and pronouns in German and Hebrew 4- and 5-year-olds.

Fig. 8. Pronouns produced in the reflexive condition in German and Hebrew.

8 Interestingly, as all the participants were tested for pronoun comprehension with two additional conditions, we could examine this girl’s performance in the other conditions. She performed at chance on the other two pronoun comprehension conditions, indicating that she did not master pronoun comprehension yet, hence explaining her use of pronouns in production.
4. Discussion

Our results indicate a very clear symmetry between comprehension and production of pronouns in language acquisition. When children allow for local coreference of the pronoun in transitive sentences in comprehension, they do so in production as well, as the Hebrew data show. When children do not allow for local coreference in comprehension, as in German, these errors do not occur in production either. Thus, what has been reported as an example of asymmetry is in fact a clear case of symmetrical development.

All the German children who were older than 4;3 showed mastery of the interpretation and production of both reflexives and pronouns. Each of the children performed flawlessly or with only a single error in the sentence-picture matching task with the pronouns, and they did not use pronouns to describe situations that targeted a reflexive in the elicitation task. Hence, the performance of the German-speaking children with respect to pronoun comprehension and production was symmetrically good. Hebrew children also showed a symmetrical performance in comprehension and production, but in this case, the performance was non-adult-like. In marked contrast to the German children, the 4- and the 5-year-old, and even some of the 6-year-old Hebrew-speakers show the classical delay in pronoun comprehension, allowing a local antecedent for the object pronoun in the sentence-picture matching task. This delay is also symmetrically reflected in their pronoun production: Hebrew-speaking children use pronouns to describe reflexive actions, namely, produce pronouns with local antecedents.

The symmetry is also shown in an analysis at the individual child level: the analysis of 50 Hebrew children who participated in both the comprehension and production tasks indicated that the children who did not master pronoun comprehension also used pronouns to describe reflexive actions in the production task and vice versa (with one exception).

From age 4, reflexives in both languages yielded good comprehension, and when reflexives were used, they were used correctly to refer to the local antecedent.

A study that tested comprehension and production of Dutch pronouns by Spenader et al. (2009) replicated earlier results on comprehension in Dutch, showing that 4 to 6 year old children sometimes (68%) interpret pronouns locally. The same study also examined production in the same children, with a picture description task. Although they claimed for asymmetry between comprehension and production, their data also showed that children who do not yet master the

---

9 Three-year-old German children do make mistakes in the interpretation of pronouns, but most of them make only few errors, and their comprehension of reflexives is not adult-like either at this age, so we cannot speak of a typical delay of principle B effect in German at this age either.
comprehension of pronouns sometimes used pronouns in the reflexive condition, although to a lesser extent than the Hebrew-speaking children in our study, namely almost 10% of the time. This lower rate could be ascribed to the slightly older age of their participants and to the difference in the tasks used. Our elicitation task was more restricted than the picture description task used in their study.

Recently, there have been discussions on the type and structure of the task that should be used to test children’s ability to comprehend pronouns (e.g., Conroy et al., 2008; Spenader et al., 2009). It has been suggested that a truth value judgment task reflects the child’s knowledge of the relevant grammatical principles, whereas the results of a picture selection task can, strictly speaking, only be interpreted as an indication of the child’s preferred interpretation (e.g., Crain and Thornton, 1998). Note that these considerations hold for production as well. A child’s production, regardless of the task used, always reflects choice and hence, a preference between possible productions. For example, a child can prefer to avoid a certain element or structure, even if the knowledge is there, or to use one of two possible elements that are both allowed by her grammar. Therefore, to test for asymmetry between production and comprehension it is better to use a comprehension test that is comparable to the production test, in that it also reflects preference. In this respect, the picture selection task that we have used, in which the child is given a certain sentence and has to indicate which interpretation is preferred, might be better suited to test for symmetry between comprehension and production.

Spenader et al. (2009) tested comprehension using a truth value judgment task, in which the children had to judge whether a spoken sentence was correctly depicted in a single picture. Thus, another explanation for the apparent asymmetry in their data, more specifically for the fact that comprehension is less target-like than production, is caused by the task that has been used. The Dutch children tested in Spenader et al. (2009) performed 68% correct on the condition that was similar to ours in comprehension, and 86% correct on the parallel condition (i.e., reflexive) condition in production, hence arguing for an asymmetry. It has been shown, though, that truth value judgment may result in about 20% lower performance than picture selection (e.g., Baauw and Zuckerman, 2009), which would mean that in fact even the results in Spenader et al. point at a symmetry rather than an asymmetry.

Thus, two findings need a theoretical explanation – one is the symmetry between comprehension and production, the other is the asymmetry between German and Hebrew. The symmetry between comprehension and production can be naturally explained on the basis of Reinhart’s (1983) suggestion that there are two types of referential dependencies that can be
established between a pronoun and its antecedent. One type of dependency is *variable binding* (governed by the binding principles), the other is *coreference* (governed by independent rules, e.g., discourse dependent rules, see Reuland, 2001). In simple transitive sentences such as ‘The fairy and the girl met, and then the fairy pinched her’, variable binding between the pronoun ‘her’ and the local subject ‘the fairy’ is ruled out by Principle B, stating that a pronoun is free in its governing category. What then needs to be checked is whether coreference is possible between ‘her’ and ‘the fairy’. Grodzinsky and Reinhart (1993) suggest that coreference is only allowed if the coreference interpretation results in a representation that can be distinguished from the variable binding representation (Rule I for Intrasentential Coreference: NP A cannot co-refer with NP B if replacing A with C, C a variable A-bound by B, yields an indistinguishable interpretation). Non-adult-like comprehension in children is then, as Grodzinsky and Reinhart (1993) argue, the result of children’s limited processing capacity, which does not allow them to compute and compare the two representations, and hence children cannot execute Rule I. This results in their inability to rule out coreference, leading to the acceptance of a local antecedent for pronouns.

The same processing limitation could also result in a failure to rule out coreference in *production*, and hence, lead to the use of a pronoun for a reflexive action, even in contexts that do not differentiate the coreferential from the variable binding interpretation. The process leading to such incorrect production would be the following: to establish referential dependencies we have two elements available: reflexives and pronouns. Now assume that the child has limited resources. That would mean that using the pronoun for a local antecedent is not ruled out immediately by applying Rule I, and hence the child has two elements available to refer to the local antecedent: reflexives and pronouns, and each of them can be chosen.

Note that this line of thought would also be helpful in accounting for the absence of the typical Principle B effect in Romance and German-speaking children and for the asymmetry between German and Hebrew. For these languages, it has been argued that accidental coreference is not (or only exceptionally) available (albeit for different reasons, see Avrutin, 1994; Baauw, 2002; Ruigendijk, 2008 for more details). That would mean that there is no possibility to choose the local antecedent for the pronoun. Since the comparison of two interpretations is not necessary to begin with, an adult-like comprehension and production is expected from a very young age. The difference found in the current study between German and Hebrew can be readily explained on this basis – German has less option for a local coreference interpretation available, hence children do not have to compute coreference against binding, and hence their limited computation does not interfere, and they choose the only option they have, both in
comprehension and in production: binding, which does not allow for local antecedents for pronouns. In Hebrew, coreference is an option, and hence, while the computation is still limited, it can be chosen for comprehension and production.

The symmetry between the comprehension and production of pronouns and reflexives is of course not expected to be limited to Hebrew and German. That is, languages with comparable pronominal systems are expected to show the same pattern. Specifically, when tested with the same material, Dutch and English children are expected to show the same pattern as the Hebrew-speaking children, namely non target-like comprehension AND production for pronouns until children are approximately 6 years old. Namely, we would expect children who accept a reflexive interpretation for pronouns to also use pronouns to describe reflexive actions. In languages in which children do not accept a local antecedent for the pronoun, such as Romance languages with object clitics, children are also expected not to use pronouns with local antecedents in production.

To summarize, this study demonstrated a clear asymmetry between languages and between pronominal elements, but not between comprehension and production, where a clear symmetry was evinced at both the group and the individual level.

References
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, 225–295.


### Table 2. Results of the production tests in German: percentage responses of the various types per each age

<table>
<thead>
<tr>
<th>Age</th>
<th>Reflexive</th>
<th>Pronoun</th>
<th>DP</th>
<th>Omission</th>
<th>No response</th>
<th>Other description</th>
<th>Reflexive</th>
<th>Pronoun</th>
<th>DP</th>
<th>Omission</th>
<th>No response</th>
<th>Other description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>49.0</td>
<td>1.0</td>
<td>8.3</td>
<td>9.4</td>
<td>3.1</td>
<td>29.2</td>
<td>39.6</td>
<td>8.3</td>
<td>3.1</td>
<td>22.9</td>
<td>2.1</td>
<td>24.0</td>
</tr>
<tr>
<td>4</td>
<td>83.0</td>
<td>1.1</td>
<td>2.3</td>
<td>0</td>
<td>0</td>
<td>13.6</td>
<td>54.5</td>
<td>4.5</td>
<td>12.5</td>
<td>9.1</td>
<td>1.1</td>
<td>18.2</td>
</tr>
<tr>
<td>5</td>
<td>83.0</td>
<td>2.3</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>13.6</td>
<td>76.1</td>
<td>0</td>
<td>9.1</td>
<td>6.8</td>
<td>0</td>
<td>8.0</td>
</tr>
<tr>
<td>6</td>
<td>92.5</td>
<td>0</td>
<td>3.8</td>
<td>0</td>
<td>0</td>
<td>3.8</td>
<td>57.5</td>
<td>5.0</td>
<td>23.8</td>
<td>10.0</td>
<td>0</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Target responses appear in bold. The shaded column is the most relevant type of error – using a pronoun with a local antecedent.
Table 4. Results of the production tests in Hebrew: percentage responses of the various types per each age

<table>
<thead>
<tr>
<th>Age</th>
<th>Reflexive</th>
<th>Pronoun</th>
<th>DP</th>
<th>Omission</th>
<th>No response</th>
<th>Other description</th>
<th>Reflexive verb</th>
<th>Pronoun</th>
<th>Reflexive</th>
<th>DP</th>
<th>Omission</th>
<th>No response</th>
<th>Other description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>15</td>
<td>26</td>
<td>0</td>
<td>15</td>
<td>17</td>
<td>24</td>
<td>4</td>
<td>69.6</td>
<td>0.0</td>
<td>7.1</td>
<td>3.6</td>
<td>7.1</td>
<td>12.5</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>31</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>12</td>
<td>2</td>
<td>69.6</td>
<td>0.0</td>
<td>3.6</td>
<td>12.5</td>
<td>1.8</td>
<td>12.5</td>
</tr>
<tr>
<td>4</td>
<td>80</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>86.7</td>
<td>1.7</td>
<td>4.2</td>
<td>1.7</td>
<td>1.7</td>
<td>5.0</td>
</tr>
<tr>
<td>5</td>
<td>69</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>89.3</td>
<td>0.0</td>
<td>4.5</td>
<td>3.6</td>
<td>0.9</td>
<td>1.8</td>
</tr>
<tr>
<td>6</td>
<td>70</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>3</td>
<td>91.3</td>
<td>0.0</td>
<td>2.5</td>
<td>3.8</td>
<td>0.0</td>
<td>2.5</td>
</tr>
<tr>
<td>Total</td>
<td>60.8</td>
<td>20.1</td>
<td>0</td>
<td>3.6</td>
<td>4.2</td>
<td>9.6</td>
<td>1.7</td>
<td>82.1</td>
<td>0.4</td>
<td>4.2</td>
<td>5.2</td>
<td>1.9</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Target responses appear in bold. The shaded column is the most relevant type of error – using a pronoun with a local antecedent.
Appendix

Verbs used in the comprehension and production tasks

<table>
<thead>
<tr>
<th>Hebrew comprehension</th>
<th>German comprehension</th>
<th>German and Hebrew production</th>
</tr>
</thead>
<tbody>
<tr>
<td>soap-mas</td>
<td>hug-mas</td>
<td>pronouns</td>
</tr>
<tr>
<td>comb-fem</td>
<td>tickle-fem</td>
<td>swing-mas</td>
</tr>
<tr>
<td>dry-mas</td>
<td>wash-fem</td>
<td>dry-fem (towels)</td>
</tr>
<tr>
<td>cover-fem</td>
<td>pet-mas</td>
<td>raise-mas</td>
</tr>
<tr>
<td>lick-fem</td>
<td>take a picture-fem</td>
<td>blind (throw light to his eyes)-mas</td>
</tr>
<tr>
<td>hug-mas</td>
<td>put oil on someone-mas</td>
<td>stab-mas</td>
</tr>
<tr>
<td>wet-mas</td>
<td>put-makeup-fem</td>
<td>wash-fem</td>
</tr>
<tr>
<td>clean-mas</td>
<td>dress-mas</td>
<td>hug-mas</td>
</tr>
<tr>
<td>wash-mas</td>
<td></td>
<td>pet-mas</td>
</tr>
<tr>
<td>pinch-mas</td>
<td></td>
<td>reflexives</td>
</tr>
<tr>
<td>wet-mas</td>
<td></td>
<td>lick-fem</td>
</tr>
<tr>
<td>cover-mas</td>
<td></td>
<td>wash-fem / soap-fem</td>
</tr>
<tr>
<td>soap-fem</td>
<td></td>
<td>hug-mas</td>
</tr>
<tr>
<td>comb-mas</td>
<td></td>
<td>cover-mas</td>
</tr>
<tr>
<td>shave-mas</td>
<td></td>
<td>pet-mas</td>
</tr>
<tr>
<td>scratch-mas</td>
<td></td>
<td>dry-mas (towels)/ cut-hair-mas</td>
</tr>
<tr>
<td>pinch-fem</td>
<td></td>
<td>put-makeup-fem</td>
</tr>
<tr>
<td>cut-hair-mas</td>
<td></td>
<td>clean-mas</td>
</tr>
</tbody>
</table>