Is the Dutch Delay of Principle B Effect dependent on verb type?

Saskia van den Akker, John Hoeks, Jennifer Spenader & Petra Hendriks

University of Groningen

Abstract

Dutch children interpret reflexives correctly from age 3 or 4 on, but frequently misinterpret object pronouns as coreferring with the local subject until age 6. We investigated whether this so-called Delay of Principle B Effect (DPBE) differs by verb type. We tested 47 children between 4 and 6 years old with regular transitive verbs (e.g., *to hit*) and grooming verbs (e.g, *to wash*), verbs that often refer to reflexive actions. In general, children displaying the DPBE performed equally well on both verb types. In contrast, children who performed poorly on reflexives as well as pronouns made significantly more errors interpreting pronouns with grooming verbs than with transitive verbs. This suggests that even young children are aware of the tendency for certain events to be self-directed. However, our results show they only apply this information when interpreting pronouns, indicating that they also use their grammatical knowledge.

1. Introduction

Until the relatively late age of 6 or 7 (Chien & Wexler 1990), English-speaking children allow an object pronoun to be coreferential with the subject of the same clause. For example, in (1) they allow the pronoun *him* to refer to *the turtle*, thus interpreting *him* identically to the reflexive *himself in* (2).

    (1)   The turtle washes him.
    (2)   The turtle washes himself.

This is surprising, because children accurately interpret reflexives from age 3 or 4. Object pronouns and reflexives are often in complementary distribution, so why are pronouns mastered years after reflexives?

This delay in pronoun acquisition also occurs in other languages, such as Dutch (cf. Koster 1993; Philip & Coopmans 1996; van Rij, van Rijn & Hendriks 2010), though recent work by Ruigendijk et al. (2010) failed to find it with German children. This is unexpected, given that German and Dutch are closely related languages.

Every pronoun acquisition study tends to choose different verbs for testing, but very little attention has been paid to semantic properties of the verbs used that might influence results. Because such studies always compare pronoun interpretation to reflexive interpretation, the set of verbs chosen includes many verbs that are relatively more likely to be used to describe a self-directed event than verbs in general: this is necessary in order to choose reflexive verbs that are natural and known to children.

The relative frequency with which an event is likely to be self-directed or other-directed is particularly salient for Dutch speakers, as in Dutch this distinction is actually grammatically marked. Dutch differs from English in that it has two reflexive forms: the simplex expression (SE) *zich* and a more complex reflexive (SELF) form *zichzelf*. These two reflexive forms tend to occur with semantically different verb types: verbs referring to events that tend to be self-directed prefer *zich,* while verbs referring to actions that are seldom self-directed are marked with *zichzelf* when reflexive. In this study, we investigate whether children's interpretation of pronouns is sensitive to these different verb types.

This paper is organized as follows. Section 2 discusses the Delay of Principle B Effect in more detail. An overview of the relation between reflexive choice and verb type is given in Section 3. Section 4 presents three possible hypotheses about the relation between pronoun interpretation and verb type. These hypotheses are tested in an experiment with Dutch children between 4 and 6 years old, which is described in Section 5. Section 6

presents the results of this experiment and Section 7 discusses the implications of our results.

2. The Delay of Principle B Effect

The Delay of Principle B Effect refers to children's non-adult-like use of Principle B of Binding Theory. Binding Theory is a module of grammar governing the use and interpretation of reflexives, pronouns and R-expressions. Relevant for the current discussion are Principle A and Principle B of Binding Theory:

    (3)   Binding Theory (Chomsky 1981)

        Principle A: A reflexive must be bound within its governing category.

        Principle B: A pronoun may not be bound within its governing category.

According to Principle A, the reflexive in (1) must be bound by the subject and cannot be coreferential with any other referent. Principle B states that the pronoun in (2) cannot be bound by the subject and must receive its interpretation through coreference with some other referent in the discourse.

Because children's interpretation of pronouns is delayed in comparison to their interpretation of reflexives, their errors with pronouns are generally referred to as the Delay of Principle B Effect (DBPE). The DBPE has been attested in languages such as English,

Dutch and Russian, but has been argued to be absent in simple transitive sentences in several other languages, including German (Ruigendijk et al. 2010).

Several explanations have been put forward for the DPBE in English and other languages. Reinhart (2006) argues that the interpretation of anaphoric pronouns requires that an interface operation is applied which she calls reference-set computation. This interface operation is a strategy of repair that is triggered by the use of an uneconomical procedure such as coreference. Reference-set computation involves constructing a reference set that includes alternative derivations. The listener then checks if the reference set contains a derivation-interpretation pair which is derived without the uneconomical procedure of coreference but instead with the more economical procedure of variable binding, while still having the same interpretation. If there is such a pair, the uneconomical derivation involving coreference is blocked. The consequence is that in normal situations *him* cannot corefer with *the turtle* in "The turtle washes him" if the resulting interpretation is the same as that of "The turtle washes himself". Because children still lack sufficient working memory capacity to apply the costly operation of reference-set computation, Reinhart argues that they resort to the strategy of guessing.

An alternative explanation, proposed by Hendriks and Spenader (2005/6), proceeds from the constraint-based framework of Optimality Theory and assumes that Principle A, but not Principle B, is a constraint of the grammar. Because the grammar does not restrict the interpretation of pronouns, pronouns are ambiguous. Therefore, pronouns may refer to

5

the local subject as well as to some other referent. To block the pronoun from coreferring with the subject, listeners have to take into account the perspective of the speaker. In the framework of Optimality Theory, this is modeled by bidirectional optimization. Because a speaker would use a reflexive rather than a pronoun for expressing coreference with the subject, coreference with the subject is not a possible interpretation of a pronoun for listeners who take into account the speaker's choices. In contrast to adults, children often fail to consider their own perspective as a listener and in addition consider the perspective of the speaker. Van Rij et al. (2010) argue on the basis of computational simulations and an experiment testing the predictions of their simulations that children's inability to optimize bidirectionally is caused by insufficient processing speed due to lack of linguistic experience.

Whichever explanation is correct, both accounts explain the interpretation of pronouns from the unavailability of a reflexive with the same interpretation. This suggests that the acquisition of pronouns may be dependent on the distribution of reflexives in the language.

3. Reflexive choice and verb type

Dutch, unlike English, has two reflexive pronouns, *zich* and *zichzelf*, both unspecified for gender and number. According to the Dutch grammar ANS (Haeseryn et al. 1997), some verbs only occur with the SE form and not with the SELF form or a non-reflexive object

(e.g., *schamen* 'to be ashamed' in (4a)), other verbs require the SELF form or a non-reflexive object and cannot occur with the SE form (e.g., *slaan* 'to hit' in (4b)), whereas a third group of verbs can occur with both reflexive forms as well as a non-reflexive object (e.g., *wassen* 'to wash' in (4c)).

(4) a.   Jan schaamt zich/*zichzelf   (inherently reflexive verb)

John is.ashamed SE/SELF

'John is ashamed'

b.   Jan slaat *zich/zichzelf   (transitive verb)

John hits SE/SELF

'John hits himself'

c.   Jan wast zich/zichzelf   (grooming verb)

John washes SE/SELF

'John washes himself'

The verb *wassen* 'to wash' belongs to a type of verbs called grooming verbs. Although grooming verbs resemble transitive verbs in that both verb types also allow non-reflexive objects, in some other respects grooming verbs differ from transitive verbs. In a forced-choice task with Dutch and German adults, Hendriks, Hoeks and Spenader (2011) found that in both languages, grooming verbs pattern with inherently reflexive verbs rather than

transitive verbs. They found that both in Dutch and German, grooming verbs strongly prefer SE reflexives. In Dutch, transitive verbs strongly prefer SELF reflexives. In German, the SE reflexive has a much wider distribution than in Dutch (Oya 2010), which was confirmed by the more or less equal selection of SE and SELF for transitive verbs in German. As a result, the choice between SE and SELF reflexives, and therefore the distinction between self-directed events versus other-directed events, is much more subtle in German than in Dutch and perhaps less salient for German children acquiring their language.

4. Pronoun interpretation and verb type

In this study, we systematically compare Dutch children's interpretation of pronouns with transitive verbs and grooming verbs. Previous experiments investigating the DBPE (e.g., Ruigendijk, Vasić & Avrutin 2006 for Dutch and Ruigendijk et al. 2010 for German) often used a combination of transitive and grooming verbs. On the basis of the preceding discussion, three hypotheses with mutually exclusive predictions can be formulated regarding the outcome of our experiment:

Hypothesis 1: Children who have not yet fully acquired the interpretation of pronouns and reflexives rely on frequency information for their interpretation. Grooming verbs have a relatively greater tendency than regular transitive verbs to occur with self-directed events

(as indicated by the choice of SE rather than SELF in Dutch, see the corpus study of Smits, Hendriks & Spenader, 2007). If children are sensitive to the frequency of the event direction, and if they have not yet fully acquired the interpretation of pronouns or reflexives, they will rely on this frequency information for interpretation. Thus, they are expected to prefer other-directed events for transitive verbs with pronouns and reflexives, and prefer self-directed events for grooming verbs with pronouns and reflexives. Hence, children are predicted to make pronoun interpretation errors mainly with grooming verbs and reflexive interpretation errors mainly with transitive verbs.

Hypothesis 2: The dual-reflexive system of Dutch influences how the DPBE manifests itself. Dutch and German both prefer SE reflexives with grooming verbs, but differ in the choice of reflexives with transitive verbs. With transitive verbs, SELF reflexives are strongly preferred in Dutch, whereas SE reflexives are a common choice in German (Hendriks et al., 2011). Several accounts of the DPBE assume that the interpretation of pronouns is determined through their competition with reflexives. It is conceivable that pronouns in object position only compete with SELF reflexives and not with SE reflexives, perhaps because sentences with SE reflexives do not express a transitive meaning. If children assign an incorrect interpretation to pronouns in contexts in which a SELF reflexive would have been the preferred form for expressing this interpretation, they are predicted to make pronoun interpretation errors with transitive verbs in Dutch (where SELF is preferred) but

not with grooming verbs in Dutch or either verb type in German (where SELF is not preferred). This would explain the lack of a DPBE in German.

Hypothesis 3: The DPBE does not depend on verb type. Hence, children are predicted to be delayed with both grooming verbs and transitive verbs.

5. Experiment

5.1 Participants

At a local elementary school, we tested 47 children (26 boys and 21 girls, age range 4;2 - 6;1, mean age 5;2). They were all native speakers of Dutch with no reported learning or cognitive disorders. In addition, we tested a control group of 20 monolingual Dutch adults.

5.2 Materials

The transitive verbs in our experiment were verbs for which adults in a forced-choice task preferred a SELF reflexive more than 80% of the time. The six transitive verbs were: *knijpen* 'to pinch', *slaan* 'to hit', *aanraken* 'to touch', *kietelen* 'to tickle', *tekenen* 'to draw´ and *fotograferen* 'to take pictures'. The grooming verbs in our experiment were verbs for which adults in a forced-choice task preferred a SE reflexive more than half of the time. The six grooming verbs were: *scheren* 'to shave', *afdrogen* 'to dry', *wassen* 'to wash',

insmeren 'rub in', *schminken* 'to make up' and *kammen* 'to comb'. Each verb was used twice in the pronoun condition, with two different referents.

Items consisted of an introductory sentence, in which the incorrect antecedent and the correct antecedent were introduced in that order in a conjunction, a test sentence with a pronoun or a reflexive, and a picture. We chose to use an introductory sentence in which the two referents are introduced in a conjunction, because we did not want pronoun interpretation to be influenced by discourse biases such as topicality (see Spenader, Smits & Hendriks 2009 for an experiment showing strong discourse effects on the interpretation of object pronouns). Half of the pictures showed an other-directed action, and the other half showed a self-directed action. An example of a test sentence with a grooming verb is presented in (5), and examples of pictures are shown in Figure 1.
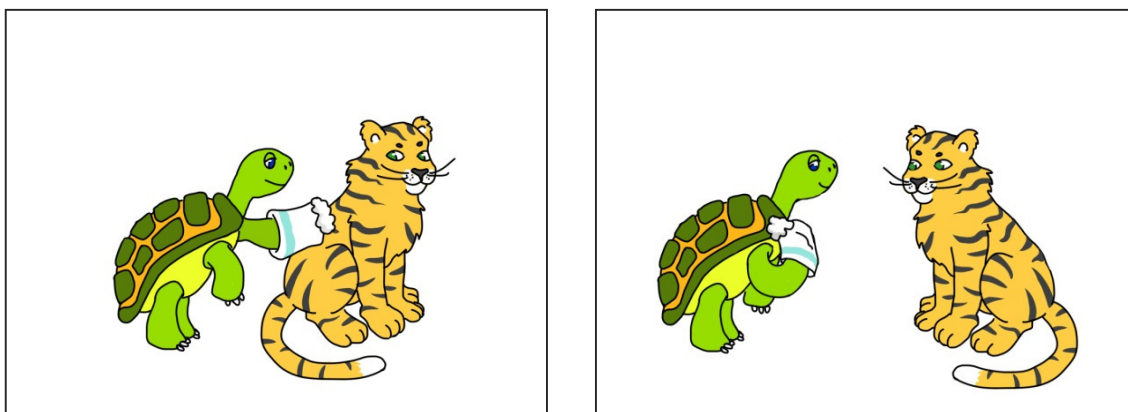

Figure 1: A picture with an other-directed action (left) and a self-directed action (right).

(5)   Hier zie je een schildpad en een tijger. De schildpad wast hem/zich met een

washand.

'Here you see a turtle and a tiger. The turtle washes him/himself with a wash

cloth.'

An example of a test sentence with a transitive verb is presented in (6), and examples of the

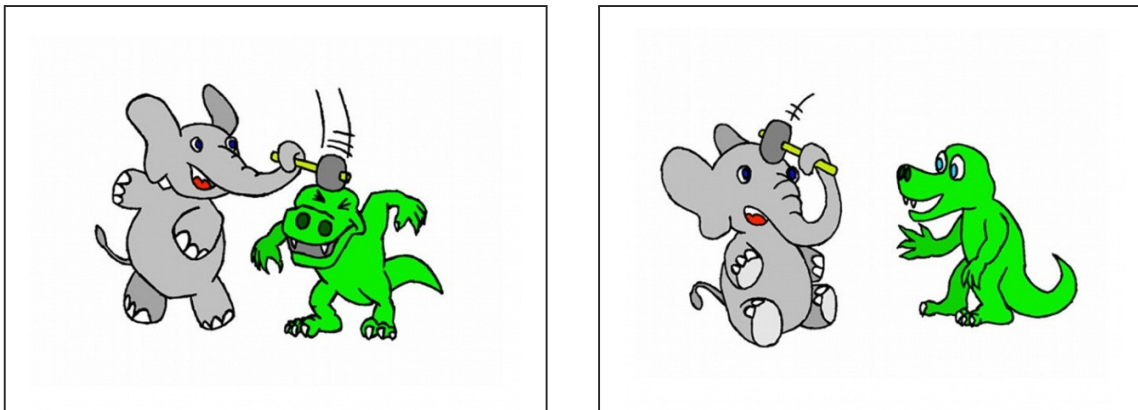corresponding pictures are shown in Figure 2.

Figure 2: A picture with an other-directed action (left) and a self-directed action (right).

(6)   Hier zie je een olifant en een dino. De olifant slaat hem/zichzelf met een

hamer.

 'Here you see an elephant and a dino. The elephant hits him/himself with a

hammer.'

The materials consisted of 12 pronoun items with grooming verbs, 12 pronoun items with transitive verbs, 6 reflexive (SE) items with grooming verbs and 6 reflexive (SELF) items with transitive verbs, making a total of 36 test items. Half of the items in each condition matched the picture and the other half were mismatch items. The pictures were balanced so that the actor would appear on the left side and the right side of the picture equally often.

5.3 Procedure

We carried out a picture verification task. First, the child was introduced to Bennie, a hand-puppet. The child was told that the hand-puppet had selected short stories for pictures with animals, but that he had made a mess of it. As a result, some stories did not match the pictures. The child was asked to help Bennie sort out which stories matched the picture and which stories did not.

The prerecorded test sentences and the pictures were presented on a laptop. During the test session, the child sat in front of the screen next to the experimenter. While the picture was shown on the screen, the introductory sentence was played on the laptop, immediately followed by the test sentence. After the child heard the test sentence, she had to decide whether the test sentence matched the picture or not. For a 'yes' response the child had to press a button with a happy green face, and for a 'no' response the child had to press a button with a sad orange face. After each 'no' response, the child was asked to explain her response. Responses and explanations were also recorded with a voice-recorder.

6. Results

First, we will give the results for all children together. We will then classify the children on the basis of their pattern of performance into four subgroups and present the results per subgroup. Following the results of the children, we will present the results of the adult control group. We conducted Repeated Measures ANOVAs on the percentages of accurate responses (arcsine-transformed), calculated per participant. Type of Anaphor (pronoun *versus* reflexive), Verb Type (transitive *versus* grooming) and Matching (match *versus* mismatch) were treated as within-participants and within-items factors.

6.1 Results of all children

There was an interaction effect between Matching and Type of Anaphor ($F(1,46)=11.93$, $p<.001$). In the match condition, the difference between pronouns and reflexives was significant (pronouns: 80%, SE=4%; reflexives: 87%, SE=4%; $F(1,46)=4.30$; $p<.05$)), but much smaller than in the mismatch condition (pronouns: 38%, SE=5%; reflexives: 64%, SE=5%; $F(1,46)=23.99$; $p<.001$). The main effect of Type of Anaphor was significant ($F(1,46)=21.05$, $p<.001$), as was the main effect of Matching ($F(1,46)=31.95$, $p<.001$). No other effects were significant.

14

6.2 Classification into subgroups

Van Rij, van Rijn and Hendriks (2010) found that the classification of children into different subgroups (based on task performance) was predictive of their linguistic behavior. We used their classification scheme to create three subgroups. Children who showed incorrect performance on both pronouns and reflexives were classified as belonging to the *Incorrect Performance group* (n=21, 10 boys and 11 girls, age range 4;2-5;8, mean age 4;9). Incorrect performance means a score on the items of less than 80% correct. Children showing correct performance on reflexives but incorrect performance on pronouns were classified as belonging to the *Delay of Principle B Effect group* (n=17, 12 boys and 5 girls, age range 4;6-6;0, mean age 5;4). The children in the third group performed correctly on both pronouns and reflexives and were classified in the *Correct Performance group* (n=7, 3 boys and 4 girls, age range 5;0-6;0, mean age 5;5). Two children (both aged 4;4) did not fall in any of these three groups, as they performed correctly on pronouns but incorrectly on reflexives. The age distribution of the three subgroups (and of the two children in the 'other' group) is shown in Figure 3.
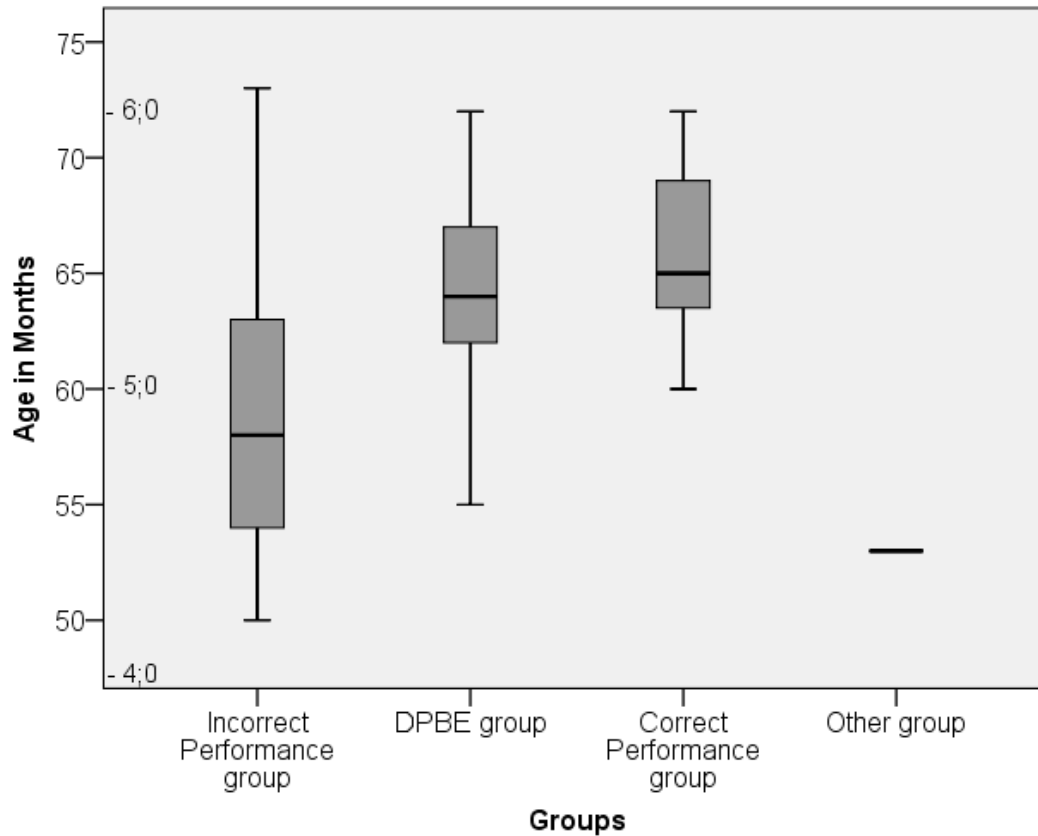
Figure 3: Mean age in months (and mean age in years on the inside of the Y-axis) for the three subgroups (and for the two children in the 'other' group) based on their performance on reflexives and pronouns.

6.3 Incorrect Performance group

For the Incorrect Performance group, there are no significant interaction effects. There is a significant main effect of Matching ($F(1,20)=15.87$, $p<.001$): matching items (80%, SE=4) were more often correct than mismatch items (34%, SE=4). This effect is probably due to a

16

yes-bias (cf. Chien & Wexler, 1990). A yes-bias results in above chance performance on match items but below chance performance on mismatch items (see van Rij et al., 2010, for discussion and a computational account of the yes-bias). No other effects were significant.
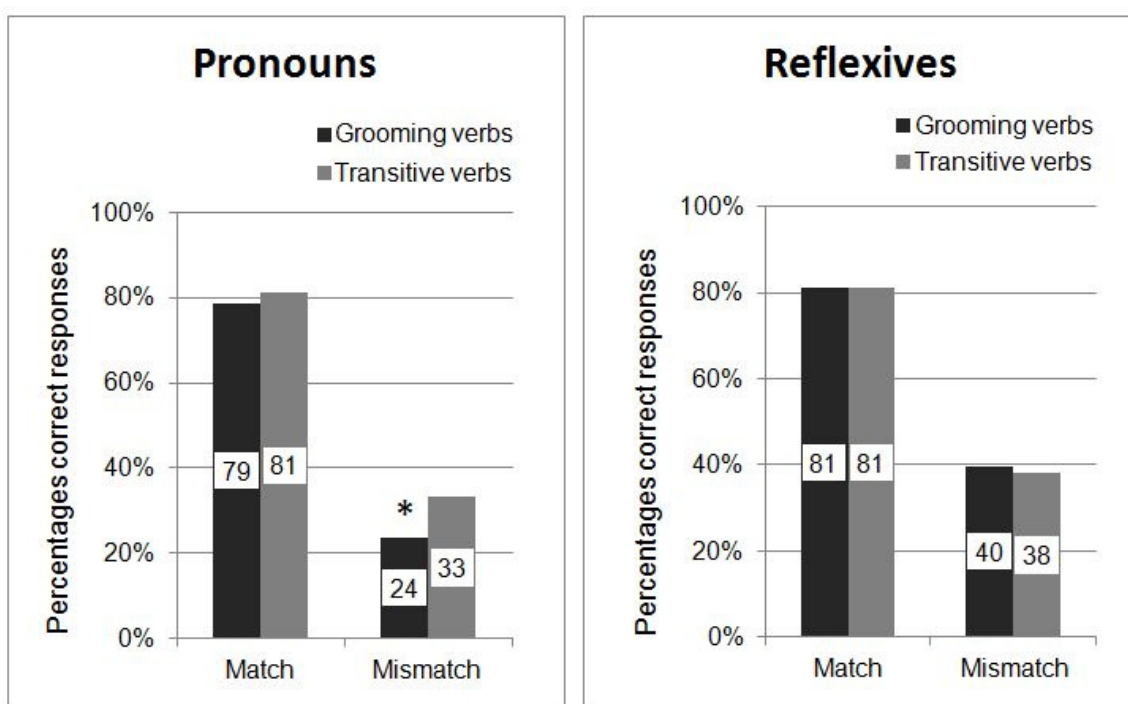

Figure 4: Accuracy results of the Incorrect Performance group (n=21); *=difference significant at .05.

We also conducted an *a priori* motivated specific comparison, based on Van Rij et al. (2010), who found an effect in the mismatch items for pronouns. To see whether this effect could be replicated in the present study, we conducted a planned comparison on the

pronoun mismatch data only. We indeed found a significant difference between the two verb types (F(1,20)=5.81, p<.05), see Figure 4. This suggests that children in the Incorrect Performance group have more difficulty interpreting mismatch items with pronouns that are objects of grooming verbs than with pronouns that are objects of transitive verbs.

6.4 Delay of Principle B Effect (DPBE) group

An interaction effect was found between Type of Anaphor and Matching (F(1,16)=36.74, p<.001). Follow up analyses showed that this interaction was caused by the difference between pronouns and reflexives being much larger in the mismatch condition (Pronouns: 24%, SE=4; Reflexives: 84%, SE=2; F(1,16)=147.27; p<.001) than in the match condition (Pronouns: 70%, SE=6; Reflexives: 96%, SE=2; F(1,16)=16.60; p<.005), see Figure 5. There were also main effects of Type of Anaphor (F(1,16)=71.215, p<0.01) and of Matching (F(1,16)=41.068, p<0.001). No other effects were significant.
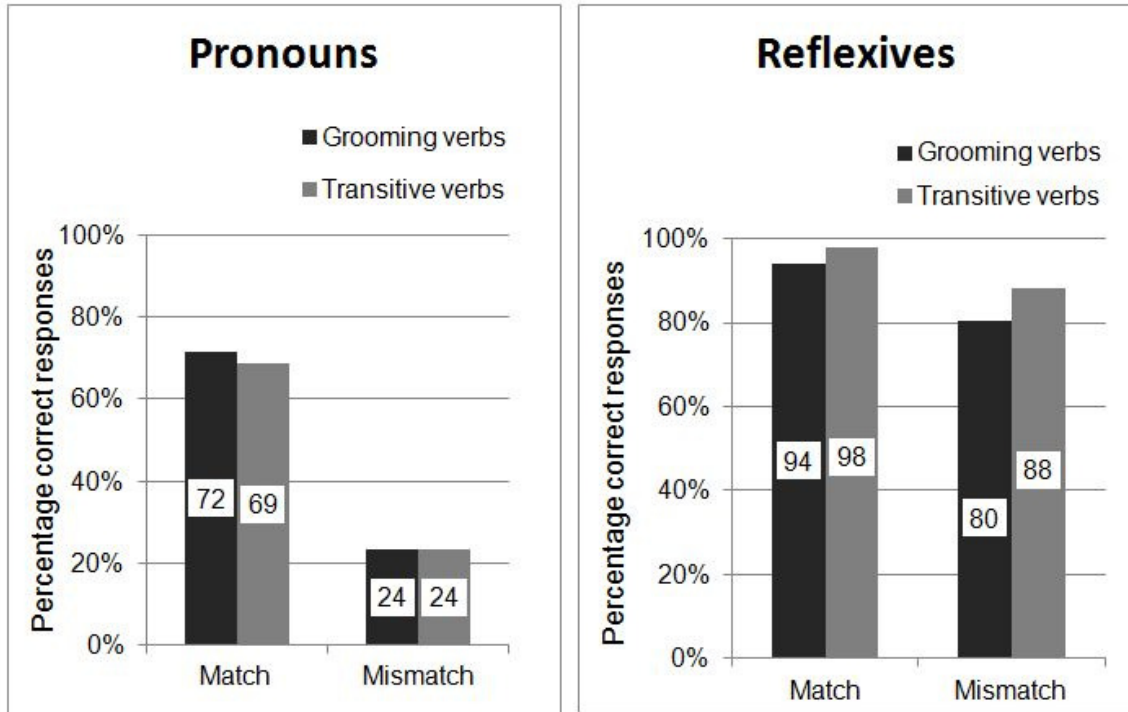
Figure 5: Accuracy results of the DPBE group (n=17).

6.5 Correct Performance group

In the Correct Performance group, a marginally significant interaction effect was found between Type of Anaphor and Matching ($F(1,6)=5.33$, $p=.06$). This interaction was due to a (marginally significant) difference between pronouns and reflexives in the mismatch condition (Pronouns: 86%, SE=4; Reflexives: 95%, SE=5; $F(1,6)=5.08$; $p=.065$), which was absent in the match condition (Pronouns: 98%, SE=2; Reflexives: 98%, SE=2; $F<1$), see Figure 6. No other effects were significant.
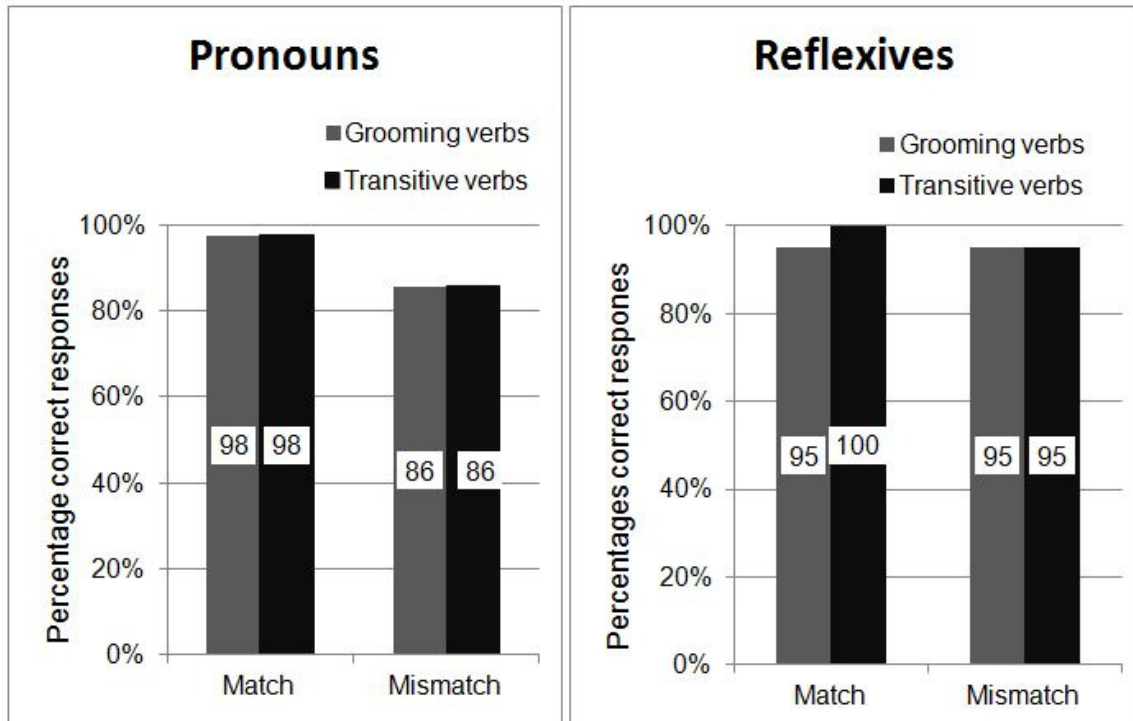
19

Figure 6: Accuracy results of the Correct Performance group (n=7).

A Repeated Measures ANOVA with Performance Group (Incorrect Performance, DPBE, Correct Performance) as a between-participants factor revealed a significant Performance Group x Type of Anaphor x Matching interaction ($F(2,40)=4.34$, $p<.05$). This indicates that the patterns that were found for each group (as described above) are indeed significantly different.

6.6 Adult control group

The adult control group consisted of 20 adults. One adult scored less than 80% correct on pronouns; these data were not included in the analysis. There was a main effect for Type of Anaphor ($F(1,18)=10.62$, $p<.005$), with adults performing better on reflexives (100%, SE=0) than on pronouns (97%, SE=1). No other effects were significant.

7. Discussion

In this study, we aimed to find out whether the DPBE is dependent on the semantic type of the verb. We compared children's interpretation of pronouns and reflexives with transitive verbs to their interpretation of pronouns and reflexives with grooming verbs. In the children who showed the DPBE pattern and performed adult-like with reflexives but made errors with pronouns, no difference in pronoun interpretation was found between the two verb types. This suggests that the DPBE is not dependent on verb type.

However, in children who not only showed incorrect performance on pronouns but also on reflexives we did find a difference between the two types of verbs. These children, who were generally younger than the children in the DPBE group, made more errors with grooming verbs than with transitive verbs in mismatch items. Our results are compatible with Sigurjónsdóttir and Coopmans (1996), who found that children make more errors with grooming verbs than with transitive verbs in Dutch complement clauses. Also, whereas

21

Norwegian children hardly ever misinterpret object pronouns with transitive verbs, they do so more often with grooming verbs (Hestvik & Philip, 1999/2000).

As grooming verbs tend to express self-directed events, whereas transitive verbs tend to express other-directed events, the frequency with which self- and other-directed events occur with pronouns and reflexives may have influenced the younger children in our study. This difference might be particularly salient for Dutch children because it is grammatically marked in the choice of reflexive form, with self-directed events being marked by SE forms and other-directed events being marked by SELF forms. So interpreting pronouns that are objects of grooming verbs may be difficult for young children because the typical self-direction of the verb conflicts with the requirement of Principle B that pronouns may not corefer with the local subject. If children have not mastered Principle B yet, their grammar allows an object pronoun to be disjoint to the local subject as well as coreferential with the local subject. The typical self-direction of the verb will then encourage them to interpret the event as a self-directed event, which results in a coreferential interpretation and hence a violation of Principle B. This does not seem to be a general effect of frequency, though, but appears to be modulated by the grammar. Hypothesis 1 stated that children rely on frequency information for their interpretation of pronouns and reflexives. However, the part in the hypothesis about reflexives was not confirmed by our results. Children did not make more errors interpreting reflexives with

transitive verbs than with grooming verbs. So frequency of the event direction only seems to affect pronoun interpretation, and not reflexive interpretation.

We did not find any evidence for Hypothesis 2, according to which the dual-reflexive system of Dutch influences how the DPBE manifests itself: the DPBE occurs just as frequently with grooming verbs as it does with transitive verbs in Dutch. This means that we cannot explain the lack of DPBE in German found by Ruigendijk et al. (2010) as deriving from the less consistent use of the SELF reflexive with transitive verbs in German. However, our finding that even the youngest children, who still made errors in reflexives, were sensitive to the preferred event direction of the verb is relevant: in Dutch, unlike German, how likely an event is to be self-directed versus other-directed is highly salient, and seems to be acquired very early. It would be interesting to test very young German children to see if they are as aware of this distinction as their Dutch counterparts. If they are not, this might explain why they do not show a DPBE for grooming verbs, which are commonly used for testing the DPBE.

References

Chien, Y.-C. & K. Wexler. 1990. "Children's knowledge of locality conditions on binding as evidence for the modularity of syntax and pragmatics". *Language Acquisition* 13. 225-295.

Chomsky, N. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Haeseryn, W., K. Romijn, G. Geerts, J. de Rooij, & M. C. van den Toorn. 1997. *Algemene Nederlandse Spraakkunst.* Groningen: Martinus Nijhoff & Deurne: Wolters Plantyn.

Hendriks, P., J. Hoeks & J. Spenader. 2011. Reflexive choice in Dutch and German. Unpublished manuscript, University of Groningen.

Hendriks, P. & J. Spenader. 2005/2006. "When production precedes comprehension: An optimization approach to the acquisition of pronouns". *Language Acquisition* 13. 319-348.

Hestvik, A. & W. Philip. 1999/2000. "Binding and coreference in Norwegian child language". *Language Acquisition* 8.171-235.

Koster, C. 1993. *Errors in anaphora acquisition.* PhD diss., Utrecht University.

Oya, T. 2010. "Three types of reflexive verbs in German". *Linguistics* 48.227-257.

Philip, W. & P. Coopmans. 1996. "The double Dutch Delay of Principle B Effect". *Proceedings of the 20th Boston University Conference on Language Development*

(BUCLD) ed. by A. Stringfellow, D. Cahana-Amitay, E. Hughes and A. Zukowski, 576-587. Somerville, MA: Cascadilla Press.

Reinhart, T. 2006. *Interface Strategies: Optimal and Costly computations*. MIT Press.

Ruigendijk, E., N. Friedmann, R. Novogrodsky & N. Balaban. 2010. "Symmetry in comprehension and production of pronouns: A comparison of German and Hebrew". *Lingua* 120.1991-2005.

Ruigendijk, E., Vasić N. & S. Avrutin. 2006. "Reference assignment: Using language breakdown to choose between theoretical approaches". *Brain and Language* 96. 302-317.

Sigurjónsdóttir, S. & P. Coopmans. 1996. "The acquisition of anaphoric relations in Dutch". *Connecting children's language and linguistic theory* ed. by W. Philip and F. Wijnen, 51-72. Amsterdam: Amsterdam Series in Child Language Development 5, University of Amsterdam.

Smits, E. J., P. Hendriks & J. Spenader. 2007. "Using very large parsed corpora and judgment data to classify verb reflexivity". *Anaphora: Analysis, Algorithms and Applications* ed. by A. Branco, 77-93. Springer-Verlag, Berlin Heidelberg.

van Rij, J., H. van Rijn & P. Hendriks. 2010. "Cognitive Architectures and Language Acquisition: A Case Study in Pronoun Comprehension". *Journal of Child Language* 37.731-766.

Spenader, J., E.-J. Smits & P. Hendriks. 2009. "Coherent discourse solves the Pronoun

Interpretation Problem". *Journal of Child Language* 36.23-52.