

# Dissolving a Dutch Delay in the Acquisition of Syntactic and Logophoric Reflexives

Peter Coopmans, Margreet Krul, Esther Planting, Ilse Vlasveld and Albert van Zoelen  
Utrecht University

## 1. Introduction

In this paper we will show that a language-specific delay in the development of Dutch children's knowledge of reflexive interpretation, which was solved as a lexical-grammatical problem, may turn out to vanish as a language-specific problem if we take a closer look at the experimental design that was used to test these children's behaviour. It specifically concerns Dutch children's comprehension of reflexives in two kinds of antecedent-reflexive environments:

- (a) binding of a reflexive element by a coargument antecedent (we will refer to this as syntactic, anaphoric binding)
- (b) binding of a reflexive by a non-coargument antecedent (we will refer to this as logophoric binding)

## 2. Background

### 2.1. Background: Part 1

In the last 10 to 15 years some robust experimental findings have been obtained indicating that children at the ages of 4 – 5 years old have problems with the proper interpretation of pronouns, while at this same age they have virtually no problems getting the correct antecedents of reflexive anaphors. A simple example to illustrate this is given in (1).

- (1) a. The girl is pointing at herself.     $\approx 85\%$  adultlike response
- b. The girl is pointing at her.        $\approx 50\%$

While sentence (1a) is only understood by these children on a locally bound reading of *herself*, sentence (1b) gives rise to some indeterminacy about whether

---

\* We thank the children and teachers of De Wantijschool at Dordrecht, De Schakel at Nieuwlekkerland, and De Ark at Zwolle. We are especially grateful to Caroline North, Bill Philip and Paulien Vrieling for their help in designing and running the experiment.

*her* can be interpreted coreferentially with *the girl* – contrary to adult judgements.

One line of explanation is based on the idea that this indeterminacy results in a kind of chance performance (hence  $\approx 50\%$ ), due to the fact that these children are still having problems determining the exact conditions under which coreference is and is not allowed. This requires assigning the same referential value to two independently referring linguistic elements, and this involves factors of a discourse-pragmatic nature, beyond grammar proper. That is why this Pronoun Interpretation Problem is sometimes referred to as a syntax-discourse interface problem, as for example in the work of Avrutin (1999), following the original insights of Chien and Wexler (1990) and Reinhart and Grodzinsky (1993). The findings for this acquisition problem in English have been corroborated for a number of other languages. But there is more.

In the acquisition of a language like Dutch it turns out that there may be an additional factor playing a role in children's comprehension of the referential properties of pronominals. In a series of experiments carried out at Utrecht University in recent years, we have found that with certain verbs Dutch children perform more poorly than with other verbs. Likewise, we have found that in certain constructions the Pronoun Interpretation Problem is more strongly visible than in others. So, for example, in constructions like (2a) we find that Dutch children show comprehension performance that we could describe as appr. 50% adultlike, no different from their English peers. But in constructions like (2b) we find even worse performance, only 25% adultlike.

- (2) a. Jan wilde dat Ad hem aaide.  $\approx 50\%$  adultlike response  
Jan wanted that Ad him patted  
'Jan wanted Ad to pat him.'  
b. Jan wilde dat Ad hem waste.  $\approx 25\%$   
Jan wanted that Ad him washed  
'Jan wanted Ad to wash him.'

In Sigurjónsdóttir and Coopmans (1996) and Philip and Coopmans (1996) it was shown that this may be due to the fact that sentence (2b) contains an inherently reflexive verb (*wassen*, "wash"), but sentence (2a) just a plain transitive one (*aaien*, "pat"). That lexical difference is a crucial factor. Such inherently reflexive verbs in Dutch require the element *zich*, and it has been argued that this is the extra source of confusion surrounding the Pronoun Interpretation Problem (see also Baauw 2000).

The generalisation that can be drawn from this research can be stated as follows: "Children accept pronominal *hem/haar* ("him/her") where adults should get *zich*." This problem does not arise for English children because English does not have a counterpart to the Dutch "se"-anaphor *zich*.

In the studies mentioned here, it was shown that the Pronoun Interpretation Problem plus additional lexical factors that cause the extra delay in Dutch can jointly be explained in terms of the Reflexivity model proposed by Reinhart and

Reuland (1993). For reasons of space, we cannot go into the details of this explanation here, but it is important that we capitalize on at least one aspect in which the Reflexivity model departs from the standard Binding Theory of Chomsky (1981): the binding principles of the Reflexivity model constrain dependencies between *coarguments*. Other anaphoric dependencies fall outside the scope of these binding principles. The explanation of such other dependencies follows from (interactions with) other principles, either grammar-internally, such as conditions on syntactic chains, or grammar-externally, in conjunction with principles of pragmatics and discourse.

## 2.2. Background: Part 2

The second piece of necessary background concerns the question how this Reflexivity model which we are assuming deals with reflexive elements, and what kind of predictions it makes with respect to children's comprehension of these. In the framework adopted it is necessary to distinguish between so-called *syntactic reflexives* and *logophoric reflexives*.

A reflexive element, like *himself*, can be bound as a syntactic reflexive if its antecedent is a coargument, that is, if the antecedent and the reflexive anaphor share the same predicate. A reflexive can also have a non-coargument as its antecedent, and then conditions on logophoric construal come into play. A well-known example is the sentence *A picture of myself would be nice on this wall*, where the sentence itself does not contain a syntactic antecedent, yet has the fully acceptable interpretation where the antecedent of the reflexive is the person whose perspective, desire or intention is at the centre of the sentence expressed.

The task for the acquisition researcher would thus require investigating children's comprehension of reflexives in the core syntactic environments regulated by reflexivity (syntactic binding) and other, non-core environments (logophoric binding). Two such environments are presented in (3) respectively.

- (3) a. The man was hiding himself.  
b. The man hid a book behind himself.

In example (3a) the object of hiding, *himself*, is a coargument of the subject *the man*, and the local relationship between these two elements is an instance of syntactic binding. In example (3b) *himself* is part of the locative PP [behind himself], and is not an argument of the verb *to hide*, hence not a coargument of *the man*. Construal between these two elements is not a form of syntactic binding, but an instance of logophoric binding. As is well-known, in such environments the complementary distribution between anaphors and pronouns breaks down. If sentence (3b) contained the pronoun *him*, this element would also be interpretable with *the man* as its antecedent.

The research on the Pronoun Interpretation Problem has shown that children's comprehension is delayed as a result of the fact that they have not mastered yet the precise discourse-pragmatic principles under which coreference

is possible. We would expect a similar effect - a similar delay in the comprehension of logophoric reflexives - where discourse-pragmatic factors determine what can and cannot be a possible antecedent. Not so for anaphoric reflexives, whose interpretation is regulated purely syntactically. Hence we can formulate our acquisition hypothesis as follows:

(4) Syntactic binding of reflexives is automatic, computation is straightforward, hence we expect early mastery. Non-syntactic binding (i.e. logophoric binding) requires access to discourse properties and constraints, which is cumbersome. The latter should be reflected in non-adultlike performance.

This research question and the corresponding hypothesis themselves are not new. The question was addressed by Avrutin and Cunningham (1997), who investigated it in an experiment with American English children. Avrutin and Cunningham subjected these children to a Truth-Value Judgement Task, in which a handpuppet commented on situations like those described in sentences (5a) and (5b).

- (5) a. The man near the boy was washing himself. (anaphoric)  
b. The man near the boy hid a book behind himself. (logophoric)

The children they investigated correctly judged (5a) to be good where *himself* referred to *the man near the boy*, and correctly disapproved of *himself* referring to *the boy* (96% correct). Sentences like (5b) turned out to be more problematic. In particular, Avrutin and Cunningham found that these children displayed 66% adultlike performance in cases where the reflexive *himself* in the locative PP incorrectly referred to *the boy* inside the complex subject *the man near the boy*.<sup>1</sup> They interpreted this finding as support for the acquisition hypothesis. The developmental path for mastering logophoric construal is harder, and it takes children a longer time to master the relevant factors.

### 2.3. Background: Part 3

Coopmans and Avrutin (1999) administered the same kind of test to Dutch children, with both the bare reflexive *zich* ("se") as well as the complex reflexive *zichzelf* ("him/her-self") in both positions: as the object of a transitive verb, and as the complement of a locative preposition. They found that the results for English were not totally replicated for Dutch. In particular, they found a surprisingly poor performance on reflexives in coargument position. (We will discuss some of these results in more detail below.) This seems to point to interlinguistic variation in the domain of reflexive binding.

---

<sup>1</sup> Avrutin and Cunningham arrived at these percentages after excluding some troublesome trials containing the verb *cover*, which gave rise to some unexpected visual errors.

The explanation that Coopmans and Avrutin proposed for this extra delay in Dutch was, again, one based on a *lexical* property of Dutch *zich* (and correspondingly *zichzelf*). They showed how lexical-grammatical features of *zich* are a source of confusion for the Dutch language-learning child, in fact the same source of confusion that lies at the heart of the extra language-specific problems Dutch children have in getting the proper interpretation of Dutch pronouns. For reasons of space we cannot go into the exact details of their proposal, but simply state here that their explanation is a theoretical account of the following combined generalisations:

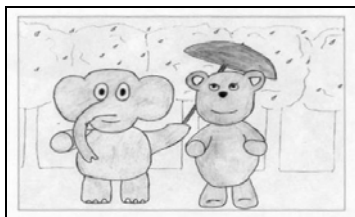
- (6) a. Children accept pronominals in adult environments of *zich*;
- b. Children accept *zich* in adult environments of pronouns.

Once again, since English does not have *zich*, English children do not have to deal with this source of extra confusion, and therefore they do not have this extra delay.

### 3. Criticism

Our point of criticism that we would like elaborate in this paper is that Coopmans and Avrutin may have jumped too hastily to the conclusion that the poor performance of Dutch children necessarily points to a lexical-grammatical account involving the features of Dutch *zich* (and correspondingly *zich-zelf*). We would like to note, though, that the position taken by them is fully consistent with earlier accounts that (extra) delays may be the combined effect of grammar-discourse interface problems and lexical-syntactic features of anaphors. Indeed, fine-grained, theoretically plausible accounts can be offered for such inter- and intra-linguistic variation. Yet, we believe that their results may have been due to flaws in the experimental set-up, and that this should be investigated seriously. Replication is needed to show that the results are not misleadingly taken to support this "lexicon & grammar & discourse" approach.

Figure 1 contains a picture of one of the Logophoric-False trials of their experiment, with the target sentence given in (7). Table 1 gives the percentages correct results on the Logophoric conditions True and False.



**Figure 1. Logophoric-False (*zich*)**

- (7) De olifant naast de beer houdt een paraplu boven zich.  
The elephant next to the bear holds an umbrella above SE

**Table 1.**

age groups	True (L)	False (L)
4;1-5;0	76%	27%
5;0-6;6	79%	27%

Let us look at these results in somewhat more detail. The handpuppet's comment on this picture is that the elephant is holding an umbrella above *zich/himself*, that is, the elephant, which is obviously incorrect (hence, Logophoric-False). Both age groups in the Dutch experiment performed very poorly on this Logophoric condition, as indicated by the 27% correct scores. That itself would still be consistent with the acquisition hypothesis, but the following example shows that things are more complicated.



**Figure 2. Anaphoric-True (zichzelf)**

- (8) Het babyaapje naast de moederaap wast zichzelf.  
Baby monkey next to Mummy monkey washes himself

**Table 2.**

age groups	True (A)	False (A)
4;1-5;0	80%	43%
5;0-6;6	87%	90%

Example (8) in conjunction with Figure 2 is an example of an Anaphoric condition, with the reflexive *zichzelf* in the object position of the verb (This one happens to be good, hence True). But if we look at Table 2 of percentage correct responses on the Anaphoric Conditions True and False, we observe poor performance (43%) in the younger age group on the Anaphoric-False condition. The older age group is doing fine on both True and False versions of the Anaphoric condition – as expected – , but the younger age group's performance is quite contrary to expectation.<sup>2</sup>

<sup>2</sup> It should be noted that the older group's performance on *zich* (rather than *zichzelf*) in the Anaphoric-False condition is also very poor: 26% correct.

The main problems that the experimental results pose for Coopmans and Avrutin can be summarised as in (9):

- (9) a. All children incorrectly allow *zich* (anaphoric and logophoric) to be linked to non-c-commanding sentence-internal antecedent.

\* [NP N [PP.. NP.. ]] V [ .. *zich* ]

- b. Young children are at chance in similar conditions for *zichzelf*. Older children do indeed perform better on coargument *zichzelf*, not logophoric *zichzelf*.

In comparison with the English data of Avrutin and Cunningham (1997), the Dutch data seem rather messy. The experimental result in (9a) suggests that Dutch children treat *zich* as if it were a regular pronominal ("him/her"), in which case the construal would be comparable to "The man next to the *boy* washes *him* / holds an umbrella above *him*". That is the empirical basis for their generalisation in (6b).

Let us look at the experimental set-up in some more detail. If we take the case of the elephant and the bear as our piece of illustration, we should remark that in that very picture use of the phrase *the elephant next to the bear* is no more informative than just *the elephant*. After all, in the picture there is just one elephant, and *next to the bear* is not really used as restrictively modifying PP. It does not contribute to setting this elephant apart from any other elephant. The subject-internal PP, in fact, was only used to introduce a second character, which for the child – given the situation – could serve as an antecedent. But introducing it this way may itself have been a source of confusion. It is mentioned in the test sentence, so there must be a reason for it to be mentioned there. The child's attention may have erroneously been focused on precisely that character, as if it meant "the elephant is holding the umbrella above the bear". Why else would it have been inserted there in the test sentence? <sup>3</sup>

A second point of criticism concerns the idea that – in a truth-value judgement task - for the child to be able to pass a negative judgement on a sentence, it would make much more sense if the corresponding *positive* judgement has been under consideration. So, in this particular case, was there any indication in the story that the elephant could have held the umbrella above himself? Was this at any point in the story considered as a possible outcome? Crain *et al* (1996) have dubbed this the condition of "Plausible Dissent", a pragmatic felicity condition on the use of the Truth-Value Judgement Task. It is impossible to meet this condition with a single, static picture. It not only applies to stories where the test sentence is a false description of the actual outcome, but

---

<sup>3</sup> Essentially a violation of the Gricean Maxim of Manner, already used as a point of criticism in the discussion of relative clause acquisition by Hamburger and Crain (1982).

also to cases where a truthful description of the actual outcome is to be judged. The opposite situation should have been under consideration as a possible outcome.

We believe that it is worth investigating whether the results come out differently under a Truth-Value Judgement design that meets these two conditions of pragmatic felicity: (i) Felicitous use of a restrictive PP inside the complex NP subject; (ii) Satisfying conditions of plausible deniability. Our reason for setting up a new version of the experiment, which we will discuss in the next section, is that "the delay may dissolve under a disciplined design".

#### 4. Experiment

We have revised the experiment (TVJ design), and investigated a new group of children's focussing primarily on their comprehension of the reflexive *zich* in anaphoric and logophoric environments. We tested 109 children (ages between 3;5-7;9), 10 of whom were excluded because they failed to pass the control conditions that were relevant for this task. We will return to one of these control conditions below. The group of 99 children was divided into a younger and an older age group, as is shown in Table 3 below.

Rather than modifying adjunct PPs, like *next to the bear*, we used PPs following relational nouns ("the father of ..."), and made sure that throughout the experiment it was clear who was "the father of ..", "the mother of .." etc. We did this by introducing one big family (the boy Flipje and his parents, the girl Rose and her parents etc.) that formed the set of protagonists in all the scenarios that the children were asked to judge (test sentences, filler sentences, control sentences etc.) One of the first control sentences was used to investigate whether the children tested could parse a complex NP like "the father of X" correctly. If they failed, they participated in the test, but their scores were excluded from the results obtained. By doing this we circumvented our first point of criticism levelled against the design of Coopmans and Avrutin (1999).

Conditions of pragmatic felicity were also met by introducing 4 characters rather than 2 in each story, so that we could felicitously talk about the father of X as opposed to the father of Y etc. We also made sure that, by having each scenario consisting of a series of pictures with the rationale for more than one activity (actual or possible), we would observe the conditions of Plausible Dissent/Assent, where denial or acceptance of the test sentence would be equally plausible as possible (not actual) outcomes of the scenario. Here we spell out one example to illustrate how we felt we could meet these criteria. All stories were accompanied by pictures in a storybook which was read to the child. The handpuppet was listening in but could not see the pictures, as he was sitting on the opposite side of the table.

(10) First Picture, containing 4 characters. (These had been introduced before, in control scenarios and in filler items.)

*Here we see Flipje and Flipje's father. And here we see Rose and Rose's father. They were all taking a walk.*

Second Picture: (4 characters, 3 umbrellas)

*Suddenly it began to rain. Flipje and Flipje's father managed to stay dry. But there was only one umbrella for Rose and Rose's father. Rose didn't want to become wet, but neither did Rose's father. Then Rose asked her father: could you hold the umbrella above me, or will you hold it above yourself?*

Third Picture: (Rose's father holding umbrella above himself)

*Kermit, can you tell us what happened then?*

And then the handpuppet responded by uttering the following sentence in Dutch, this time a correct description of the actual outcome:

*Ik weet het, de vader van Rose hield de paraplu boven zich.*

'I know, the father of Rose held the umbrella above SE.'

So, this would be Logophoric-True.

The experimental conditions we used can be divided up into the following subclasses, for both the Logophoric and the Anaphoric conditions. To create the logophoric environments, we used the Dutch counterparts of "hold above", "put down next to", and "hide behind". For the anaphoric environments, here with *zich*, we used the following predicates, whose Dutch counterparts can be either purely transitive or inherently reflexive: "wash", "dress" and "dry".

Further relevant features of the experimental design were the following:

- i. the conditions were counterbalanced in 4 sets;
- ii. each set contained 2 *zich*-stories, each story containing 2 conditions;
- iii. each story contained both a True- and a False-condition;
- iv. each story contained both a Logophoric and an Anaphoric condition;
- v. hence, all 4 conditions (Anaphoric-False, Anaphoric-True, Logophoric-False, Logophoric-True) were present in each set;
- vi. no predicate was used twice in a set.

In Table 3 we present the results for both age groups of children, as well as for a control group of 26 adults.

**Table 3.**

age group	n	mean age	age range	AnF	AnT	LoT	LoF
younger	49	4;11	3;5-5;7	94% (4)	82% (6)	69% (7)	69% (7)
older	50	6;4	5;8-7;9	100%	96% (3)	90% (4)	90% (4)
all children	99	5;8	3;5-7;9	97% (2)	89% (3)	80% (4)	80% (4)
adults	26	38	18-73	100%	96% (4)	100%	100%

Standard error in parentheses. Sign (2-tailed) AnF vs LoF significant for the younger group:  $p=0.004$ ; not so for the older group:  $p=0.062$ . Significant effect of age (young vs old children,  $p=0.0111$ ; all children vs adults,  $p=0.0128$ ) for both Logophoric conditions (Kruskal Wallis, with H adjusted for ties), not so for the Anaphoric conditions.

We find that the children in the younger age group perform better on the Anaphoric-False condition (AnF) than on the Logophoric-False condition (LoF). The overall group results are 94% versus 69%, and this difference is significant. This difference disappears with the older age group, who performed better on both conditions. More generally, we find significant effects of age (younger vs older children, and all children vs the adults) for the Logophoric conditions, not so for the Anaphoric conditions. In so far as these children had problems with interpreting reflexives, the problems really involved logophoric binding, not anaphoric binding, as expected under the original hypothesis.

Indeed, the children performed better on the crucial Anaphoric vs the crucial Logophoric condition, significantly so for the younger group. We see significant progress by age in the Logophoric-False condition. The results are not as sharp as in the original Avrutin and Cunningham (1997) experiment for English, but much less messy than in Coopmans and Avrutin (1999), and thus we have obtained support for the acquisition hypothesis in (4). No less important is our observation that the extra delay in Dutch does indeed dissolve under a disciplined design.

## 5. Discussion

It is very satisfying to be able to draw these conclusions, but we would still like to point to the following questions. In this experiment, we only tested children's comprehension of anaphoric and logophoric *zich*. What about *zichzelf*? We have obtained the results on *zichzelf* by a comparable older age group, and their scores are very much similar to the ones given here for the older age group on *zich*. We intend to subject a younger group of children to these test sentences including *zichzelf* shortly. In addition, we will also focus on children's comprehension of other elements in the same environments, like the regular pronouns *hem/haar*, to see if the results parallel those on *zich*.

The immediate question arises what exactly caused the results of the original experiment on English by Avrutin and Cunningham (1997) to come out as they did. Why were the results so much cleaner, why could they be used immediately in support of the acquisition hypothesis? More specifically, why did the English children perform much more adultlike on the anaphoric - coargument - *himself* (or *herself*)? A plausible reason that suggests itself involves a linguistic difference between the ways Dutch and English children could possibly analyse *self*-anaphors in their respective languages.<sup>4</sup> The appearance of the English *self*-anaphor in the object position of a transitive verb reflects its role as a true reflexiviser in the Reflexivity model of Reinhart and Reuland (1993). That is its prime function, and syntactic binding by the coargument subject is the first grammatical option. However, a similar surface situation in Dutch is in principle compatible with an alternative analysis. The sequence *zichzelf* may also be analysed as consisting of two truly separate elements, where *zelf* is an intensifier, associated with the subject NP. So, in example (8) the *zelf*-element would be associated with the subject NP, but that could leave the *zich*-element to be incorrectly linked with the NP inside the subject, i.e. Mummy monkey. Under this analysis, whatever would explain the children's poor performance on the *zich* examples could simply be carried over the *zichzelf* examples.

The problematic aspects in the experimental set-up may have been hidden in the English case, if there was no option for the English children other than to analyse examples such as (5a) as containing a reflexiviser. We noted earlier that English children do not have the extra problem of learning the special properties of *zich*, and do not get confused by it. Neither do they have the problem of misanalysing the *self*-part in complex anaphors as an intensifier. An intensifier function in English is assigned to *himself* as a whole, not to a part of it, as is shown in (11a/b). In Dutch it is just the opposite (11c/d).

- (11) a. \*John did it self.  
b. John did it himself.  
c. Jan deed het zelf.  
d. \*Jan deed het zichzelf/hemzelf.

This lexical difference between *himself* and *zichzelf* is a real one, raising further questions about the ways these complex anaphors are acquired in English and Dutch respectively. This difference may have had an effect on the different outcomes of the earlier experiments, but we admit that at this stage it remains somewhat speculative.

At any rate, by carrying out a revised experiment we have been able to take away one language-specific delay, and we have thus been able to provide clearer and stronger evidence for the hypothesis that logophoric binding is not acquired

---

<sup>4</sup> This is an idea that we borrow from the analysis in Coopmans and Avrutin (1999).

as fast as syntactic binding. This in turn supports the general idea that these dependencies need to be captured on different levels of representation.

## References

- Avrutin, Sergey (1999) *Development of the Syntax-Discourse Interface*, Dordrecht: Kluwer Academic Publishers.
- Avrutin, Sergey and Jennifer Cunningham (1997) Children and Reflexivity, *Proceedings of BUCLD 21*, 13-23.
- Baauw, Sergio (2000) *Grammatical Features and the Acquisition of Reference. A Comparative Study of Dutch and Spanish*. Doctoral dissertation, Utrecht University.
- Chien, Yu-Chien and Kenneth Wexler (1990) Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics, *Language Acquisition 1*, 225-295.
- Chomsky, Noam (1981) *Lectures on Government and Binding*, Dordrecht: Foris Publications.
- Coopmans, Peter and Sergey Avrutin (1999) A Syntax-Discourse Perspective on the Acquisition of Reflexives in Dutch, paper presented at GALA, Potsdam.
- Crain, Stephen, Rosalind Thornton, Carole Boster, Laura Conway, Diane Lillo-Martin and Elaine Woodams (1996) Quantification without Qualification, *Language Acquisition 5*, 83-153.
- Grodzinsky, Yosef and Tanya Reinhart (1993) The Innateness of Binding and Coreference, *Linguistic Inquiry 24*, 69-102.
- Hamburger, Henry and Stephen Crain (1982) Relative Acquisition, in Stan Kuczaj (ed) *Language Development: Syntax and Semantics*, Hillsdale, NY: Lawrence Erlbaum.
- Philip, William and Peter Coopmans (1996) The Role of Feature Specification in the Acquisition of Pronominal Anaphora in Dutch, *Proceedings of NELS 26*. 241-255.
- Reinhart, Tanya and Eric Reuland (1993) Reflexivity, *Linguistic Inquiry 24*, 657-720.
- Sigurjónsdóttir, Sigríður and Peter Coopmans (1996) The Acquisition of Anaphoric Relations in Dutch, *Amsterdam Series in Child Language Development 5*. 73-106.