LANGUAGE ACQUISITION, *8*(3), 171–235
Copyright © 1999/2000, Lawrence Erlbaum Associates, Inc.

Binding and Coreference in Norwegian Child Language

Arild Hestvik
University of Bergen

William Philip
Utrecht University

Four experimental studies were designed to test, in Norwegian, the hypothesis that children's nonadultlike interpretations of pronouns may be partly attributable to a lexical factor interacting with the A-Chain Condition (Philip and Coopmans (1996), Sigurjónsdóttir and Coopmans (1996)). Analyzing the antisubject orientation of pronouns in adult Norwegian as an A-Chain Condition effect at Logical Form, we predicted that, due to imperfect representation of the lexical features of pronouns, Norwegian children would significantly more often show nonadultlike understanding of antisubject-oriented pronouns than of pronouns whose adult interpretation is also constrained by Principle B. These predictions were confirmed, providing evidence for the independent effects of the A-Chain Condition and the Binding Principles on the distribution of pronouns and identifying lexical interactive effects with the A-Chain Condition as a potential locus of development and variation across child languages.

1. INTRODUCTION

One much-studied pronoun comprehension error frequently observed with preschool children consists in the apparent assignment of an identical semantic value to a pronoun and a noun phrase (NP) in contexts in which this is ruled out by the adult grammar. This is the so-called delay of Principle B phenomenon of child language. Such child errors are generally not taken to indicate incomplete knowledge of Principle B. Rather, they are taken to indicate the overgeneralization or dysfunction of a pragmatic- and discourse-related principle that, in special circumstances in adult language usage, licenses the coreferential interpretation of a pronoun and a counterindexed NP (Chien and Wexler (1990), Grodzinsky and

Requests for reprints should be sent to Arild Hestvik, Department of Linguistics and Comparative Literature, University of Bergen, Sydnesplassen 7, N–5007, Bergen, Norway. E-mail: hestvik@uib.no

Reinhart (1993), Wexler and Chien (1985)). To give it a pretheoretical descriptive name, we refer to this child comprehension phenomenon as the *pronoun semantic identity error*.

Recently, Sigurjónsdóttir and Coopmans (1996) and Philip and Coopmans (1996) proposed that child pronoun semantic identity errors can also result from an overgeneration of A-Chains due to incomplete lexical knowledge of the grammatical features of pronouns that interact with the A-Chain Condition. We reinterpret this proposal as the hypothesis that, due to lexical performance factors, children acquiring certain languages (e.g., Dutch, Norwegian) often fail to fully represent at Logical Form (LF) the complete lexical specification of a third-person pronoun, even though they have fully acquired knowledge of the lexical specifications of pronouns in the adult lexicon and can in principle construct fully adultlike LFs for any sentence containing a pronoun. Under this lexical performance hypothesis, a child acquiring a language of the appropriate type may sometimes assign to sentences of certain types lexically deviant but syntactically well-formed LFs in which a pronoun is bound by an adult-ungrammatical antecedent. This binding makes a reading available to the child that is not available to the adult and thereby gives rise to another kind of child pronoun semantic identity error that is completely independent from the discourse- and principle-related type mentioned earlier. In this article, we test the predictions of this hypothesis in an experimental study of Norwegian preschool children's interpretation of pronouns, focusing in particular on antisubject-oriented pronouns.

The article is organized as follows: In section 2, we discuss our theoretical assumptions concerning the grammatical constraints on intrasentential pronominal anaphora in adult and child grammar. In section 3, we describe the antisubject orientation of Norwegian pronouns, and we develop a theory in which this effect is derived from the A-Chain Condition applying at LF. In this section we also discuss the predictions our analysis of adult Norwegian makes for child Norwegian, when combined with the lexical performance hypothesis that children are frequently unable to fully integrate the grammatical lexical content of pronouns into their representations of sentences containing pronouns. In sections 4 through 8 we present and discuss four experiments with Norwegian preschool children. Finally, in section 9 we summarize our findings and discuss their significance for linguistic theory and for research on first-language acquisition and development.

2. THEORETICAL BACKGROUND

2.1. Principle B and the A-Chain Condition

Following Reinhart (1983), we assume that the semantic value of a pronoun can be determined in one of two ways: either syntactically via binding (*c*-command and coindexation), with subsequent obligatory semantic interpretation as a bound

variable, or nonsyntactically via reference assignment, which is not mediated by syntactic indexes.[1] Each mode of fixing the semantic value of a pronoun is constrained by principles applying at the syntax–semantics interface and the semantics–discourse interface, respectively. The principles governing syntactic binding are Principles A and B of the Binding Theory and the A-Chain Condition. We adopt the formulation of these principles in the "reflexivity" framework of Reinhart and Reuland (1993). The semantics–discourse interface principle constraining coreference in the case of nonsyntactic pronoun reference assignment is a rule that compares different linguistic representations of a string to meaning representations in a given context of use. We adopt Reinhart's "Rule I" description of this mechanism, as formulated in Grodzinsky and Reinhart (1993).

Reinhart and Reuland (1993) accounted for syntactic binding in terms of the interaction of two independent modules. These are Principles A and B of the Binding Theory, on one hand, and the A-Chain Condition, on the other hand—each of which may be described as follows (abstracting away from properties irrelevant to our proposal):

Binding Theory:
(i) Principle A: A reflexive marked predicate must be semantically reflexive.
(ii) Principle B: A semantically reflexive predicate must be reflexive marked.

A-Chain Condition:
The tail of an A-Chain must be [–R].

Focusing on Principle B, a predicate is semantically reflexive if two of its coarguments have identical semantic values.[2] For example, for a sentence expressing the proposition that John likes himself, Principle B would require that one of the syntactic arguments of the predicate carry reflexive morphology (or that the verb be inherently reflexive). Thus, the English surface form mapping to the LF interpreted as John(λx(x likes x)) can only be *John likes himself*; it cannot be *John likes him*.

Independently of the Binding Principles, the A-Chain Condition also constrains the distribution and indexing of pronouns. A-Chains are automatically

[1]Alternatives exist. For example, Berman and Hestvik (1997) developed a theory of pronominal reference resolution that does not draw the distinction between syntactic binding and discourse reference assignment. Note that this does not necessarily entail that such a theory cannot make distinctions evidenced by child development grammars, as long as some distinction can be found in the theoretical representation of the two types of constructions. Further discussion is outside the scope of this article, however.

[2]For the purposes of this article, a semantically reflexive predicate may simply be thought of as a predicate having coindexed coarguments at LF. However, as Reinhart and Reuland (1993) pointed out, to capture other effects of Principle B, such as *[The queen$_1$ invited both [Max and her$_1$]$_2$ to our party], additional assumptions are needed to accurately define semantically reflexive predicates.

formed between coindexed and subjacent[3] elements in A-positions (cf. Reinhart and Reuland (1993, 693)). To satisfy the A-Chain Condition, the tail (i.e., the *c*-commanded element in the chain) of a well-formed A-Chain must be underspecified in ϕ- and structural Case features (Reinhart and Reuland (1993, 697)). Adopting Reinhart and Reuland's terminology, we label [–R] those anaphoric elements that are sufficiently underspecified to tail an A-Chain and [+R] those that are not.[4] For example, a trace of movement is [–R]. The English *himself* is labelled [–R] because the head *-self* lacks a distinction between nominative and accusative. Hence, it is sufficiently "Case-deficient" to tail an A-Chain, and it does so in grammatical sentences such as *John hit himself* (we label such A-Chains as *interpretive A-Chains* to distinguish them from chains formed by movement). Similarly, Norwegian reflexives *seg* and *seg selv*, both of which can mean either 'himself', 'herself', 'itself', or 'themselves', are also [–R] and can tail interpretive A-Chains. In contrast, both English *him* and Norwegian *ham* 'him' are [+R] and therefore cannot be the tail of an interpretive A-Chain. Thus, in a sentence such as *John hit him*, the A-Chain Condition rules out coindexation of *John* and *him* because this would give rise to an interpretive A-Chain tailed by a [+R] expression.

Given this dissociation of the A-Chain Condition and Principle B, three logically distinct types of bound pronoun are distinguishable: (a) Antecedent and pronoun are coarguments and form an A-Chain (we call this *local binding*); (b) antecedent and pronoun are not coarguments but are still subjacent so that an A-Chain is formed; and (c) antecedent and pronoun are not coarguments, and furthermore not subjacent, so no A-Chain is formed (we call both of the latter cases *nonlocal binding*). The three logical possibilities are illustrated in (1):

(1) a. *John$_1$ likes him$_1$.
    b. *John$_1$ expects [him$_1$ to win].
    c. John$_1$ thinks that [he$_1$ will win].

The LF in (1a) is ungrammatical because both Principle B and the A-Chain Condition are violated. The LF in (1b) is ungrammatical because the A-Chain Condition is violated; however, Principle B is vacuously satisfied in this LF because [John$_1$] and [him$_1$] are not coarguments. The LF in (1c) is grammatical because both the A-Chain Condition and Principle B are satisfied. Separated by a tensed

[3]Two elements α and β are subjacent if they stand in an antecedent government relation; that is, α m-commands β, and no barrier intervenes between α and β (cf. Chomsky (1986)).

[4]Note that being [–R] does not preclude the possibility of independent reference. Dutch first- and second-person pronouns are [–R] in that they may tail A-Chains, but they also may have discourse antecedents. On this and related issues, see also Safir (1997), who argued that the notion of A-Chain used in this article is a spurious entity and that its effect should instead follow from another version of Principle B. We do not investigate the consequences of Safir's proposal on the interpretation of our experimental results.

clause boundary, [John$_1$] and [him$_1$] are not subjacent; therefore, they cannot be related by an interpretive A-Chain, and the A-Chain Condition is vacuously satisfied. In addition, [John$_1$] and [him$_1$] are not coarguments, so Principle B is vacuously satisfied as well. Thus, the binding relation in (1c) is licensed by syntactic constraints.

Note that a central feature of the Reinhart and Reuland (1993) binding theory is that Principle B only regulates local binding (coargument relations), whereas the A-Chain Condition regulates both local binding and certain cases of nonlocal binding (noncoargument relations). In particular, only the A-Chain Condition is responsible for blocking binding in (1b). This feature of the theory is crucial in deriving our central experimental predictions.

Principle B and the A-Chain Condition are purely syntactic constraints, but they indirectly constrain the interpretation of pronouns as follows: If a pronoun is bound by some element, then it will necessarily be interpreted as semantically identical with this element, by virtue of being semantically interpreted as a bound variable. On the other hand, countercoindexation does not entail semantic nonidentity. Rather, the semantic identity or nonidentity of two independently referring expressions is regulated by Rule I.

2.2. Rule I

If a pronoun is not syntactically bound, it is syntactically free. Following Reinhart (1983) and Grodzinsky and Reinhart (1993, 79), we assume that the interpretation of a syntactically free pronoun is constrained by the Interpretation Rule (Rule I) operating at the LF–discourse interface:

Rule I: NP α cannot corefer with NP β if replacing α with γ, γ a variable A-bound by β, yields an indistinguishable interpretation.

To illustrate the operation of this principle, consider again the surface form *John thinks he will win*, which corresponds to (1c). This time, however, let us counterindex the pronoun and the subject NP, as shown in (2):

(2) John$_1$ thinks [he$_2$ will win].

In the case of (2), Rule I normally determines that [he$_2$] cannot corefer with [John$_1$]. If [he$_2$] were replaced by a variable bound by [John$_1$], the resulting LF would express the proposition that John has the property λx*(x thinks (x will win))*. Under a coreferential reading of [John$_1$] and [he$_2$], (2) expresses the proposition that John has the property λx*(x thinks (John will win))*. Normally, these two propositions are semantically indistinguishable. Consequentially, in ordinary contexts of language use, Rule I causes counterindexation to entail noncoreference. However, consider the same sentence under the scope of the focus operator *only* and placed in a discourse context as in (3):

(3) John is not the best athlete. No one thinks he will win the race. Only John thinks he will win.

Again, there are two possible indexations of the last sentence in this discourse fragment, the LF (4a) and the LF (4b):

(4) a. Only $John_1$ thinks he_1 will win.
b. Only $John_1$ thinks he_2 will win.

We are interested in the coreferential interpretation of (4b) (i.e., the reading in which [he_2] refers to John despite being counterindexed with the NP [$John_1$]). If [he_2] were replaced by a variable bound by [$John_1$], the resulting LF would be syntactically identical to that in (4a). However, (4a) and the coreferential reading of (4b) contribute semantically in different ways to the meaning of the whole discourse. (4a) means that John is the only person *x* such that *x* thinks *x* will win, which entails that no other person *y* thinks that *y* will win. The coreferential reading of (4b), on the other hand, means that no person other than John thinks John will win and entails nothing about other people's beliefs about themselves. When embedded in the discourse fragment, it becomes apparent that only this latter meaning is supported by the discourse as a whole. Due to the fact that the interpretation of (4a) and the coreferential reading of (4b) are truth-functionally distinct, and this distinction is supported by the context in (3), Rule I allows [$John_1$] and [he_2] to have identical semantic value in (4b).

This is a typical case in which Rule I licenses the coreferential reading of two counterindexed NPs. In principle, Rule I may allow a coreferential reading of any counterindexed NPs. Even with LFs such as (5a), Rule I may license a coreferential reading of the counterindexed NPs in special discourse contexts:

(5) a. $John_1$ likes him_2.
b. Every boy_1 likes him_2.

For example, as Chien and Wexler (1990) pointed out, the sentence *John looks like him* normally does not admit a coreferential reading of *John* and *him*. However, uttered immediately after *That must be John*, this coreferential reading becomes acceptable (see also Heim (1993)). Crucially, although Rule I can neutralize the effects of Principle B when a potential local antecedent of a pronoun is a referring expression, it does not apply at all when a potential antecedent of a pronoun is a quantified NP, as in (5b).[5] This is because nothing can be coreferential with a quantified NP. Due to the fact that Rule I looks for a

[5]Heim (1993) disputed this. For reasons of space, we do not address the ramifications of Heim's analysis.

coreference relation, it cannot apply in principle to a relation involving a nonreferential potential antecedent. In other words, the only kind of semantic identity that may obtain between a quantified NP and a pronoun is that of a bound variable interpretation of the pronoun, in which case the quantifier binds the pronoun. Finally, we assume that Rule I applies whenever coreference is in principle possible. It is not "turned on" by special discourse contexts. Rather, it is always turned on, although its neutralization of Principle B effects may only be observed in special contexts.

2.3. Rule I Failure

Numerous studies using a wide variety of different experimental methodologies have shown that in experimental contexts, preschool children generally exhibit pronoun semantic identity errors on average 50% of the time with sentences such as (5a). Chien and Wexler (1990) replicated this finding with a large sample of English-speaking children and then showed that the same children exhibited pronoun semantic identity errors significantly less often with sentences such as (5b). For example, the 44 five-year-olds who participated in their study showed pronoun semantic identity errors only 16% of the time on average with sentences such as (5b), but 51% of the time on average with sentences such as (5a).[6] Proposing what has now become a standard analysis, Chien and Wexler argued that the marked asymmetry in the number of pronoun semantic identity errors elicited from children by the two types of sentences in (5) reflects the circumstance that (i) children of all ages have fully acquired Principle B (because it is innate and available from the onset of first-language acquisition), but (ii) below a certain age children lack adultlike knowledge of the pragmatic principle, which we described as Rule I (referred to as "Principle P" by Chien and Wexler). According to Chien and Wexler (259), incomplete acquisition of Rule I leads to an overgeneralization of coferential readings with counterindexed NPs in the case of sentences such as (5a). This overgeneralization cannot occur with sentences such as (5b), because coreference is not possible in principle and only Principle B regulates pronoun reference. On this view, either Rule I is learned from the input or it is innately specified but only gradually matures, in the sense of Borer and Wexler (1987).

Alternatively, Grodzinsky and Reinhart (1993) proposed that Rule I is both innately specified and fully available from the onset of first-language acquisition, but, due to cognitive limitations, children often have difficulty using it. When they attempt to apply Rule I, a processing breakdown often occurs. This break-

[6] The 16% nonadultlike performance observed with sentences like (5b) can be attributed to "experimental noise" and need not be considered indicative of pronoun semantic identity errors.

down is the material cause of child pronoun semantic identity errors, according to Grodzinsky and Reinhart.[7] The breakdown never occurs in the case of sentences such as (5b) because Rule I does not apply to such sentences and therefore cannot break down. Adopting Grodzinsky and Reinhart's basic analysis, we henceforth label this source of child pronoun semantic identity errors as *Rule I failure*.

2.4. Lexical Failure

Sigurjónsdóttir and Coopmans (1996) and Philip and Coopmans (1996) argued that Rule I failure cannot be the only source of children's nonadultlike interpretation of pronouns in experimental contexts. In an experiment comparable in design to the fourth experiment of Chien and Wexler (1990), Philip and Coopmans found that Dutch preschool children made significantly more pronoun semantic identity errors when the potential antecedent and the pronoun were not coarguments, as in Exceptional Case Marking (ECM) contexts such as (6b), than when they were coarguments, as in (6a). For example, 22 Dutch 7-year-olds (M = 7;6) showed semantic identity errors on average 45% of the time with sentences such as (6a) but 84% of the time with sentences such as (6b):[8]

(6) a. Het meisje wijst haar aan.
the girl points her at
'The girl is pointing her out.'
b. Het meisje ziet haar bellen blazen.
the girl sees her bubbles blow
'The girl sees her blow bubbles.'

[7]Grodzinsky and Reinhart (1993) also proposed that when Rule I breaks down, the child resorts to a "guessing strategy," and this is why, on average, preschoolers show pronoun semantic identity errors roughly 50% of the time with sentences such as (5a), because they randomly oscillate between coreferential and noncoreferential readings of counterindexed NPs. However, notice that failure to compute the results of a linguistic rule does not make any predictions about subsequent linguistic behavior; it would be equally compatible with failure to compute Rule I to either never or always allow coreference. Furthermore, close examination of the findings of the fourth experiment of Chien and Wexler (1990) casts some doubt on this aspect of Grodzinsky and Reinhart's proposal because children's pronoun semantic identity errors did not distribute normally over six repeated measures of an experimental condition testing sentences like (5a). This is unexpected if chance is a principal determinant. Moreover, Philip and Coopmans (1996) found that pronoun semantic identity errors attributable to a difficulty processing Rule I did not occur 50% of the time for Dutch children of all ages; younger children showed such errors more often than older children. On the other hand, independent support for Grodzinsky and Reinhart's principal claim about processing is that such errors occur roughly as often with adult agrammatic aphasics as with preschool children (Grodzinsky, Wexler, Chien, Marakovitz, and Solomon (1993), Rosen and Rosen (1995)).

[8]Philip and Coopmans (1996) also found that with the sentence type in (6b) the same 7-year-olds showed pronoun semantic identity errors significantly more often with the pronoun *haar* 'her' than with the pronoun *hem* 'him'. This finding provides particularly compelling evidence that a lexical factor underlies this sort of pronoun semantic identity error.

Consider the two logically possible LFs for (6b) shown in (7):

(7) a. *Het meisje$_1$ ziet [haar$_1$ bellen blazen].
b. Het meisje$_1$ ziet [haar$_2$ bellen blazen].

In the adult grammar, the LF (7a) is ruled out by the A-Chain Condition, and a coreferential reading of the counterindexed NPs in the grammatical LF (7b) is ruled out by Rule I (given the discourse context in which sentences like (6) were presented in Philip and Coopman's experiments).

In child linguistic performance, Rule I failure may "let in" a coreference reading of (7b) some of the time. In (7a), the antecedent and pronoun are coindexed, so Rule I failure cannot occur (because Rule I does not apply), but this LF is ungrammatical so the children should never assign it to a sentence such as (6b). However, note that the only thing that makes (7a) ungrammatical in the adult grammar is the feature specification of the pronoun *haar* 'her', which is too "rich" for *haar* to tail a well-formed interpretive A-Chain. If a child did not know that *haar* was [+R] in the adult lexicon, he or she could assign the LF in (8a) to the sentence in (6b), in which the nonadultlike lexical property of the representation is indicated by the label [–R]:

(8) a. Het meisje$_1$ ziet [haar[–R]$_1$ bellen blazen].
b. Het meisje$_1$ ziet [zich$_1$ bellen blazen].
'The girl sees herself blow bubbles.'

The LF in (8a) is syntactically grammatical; it is only lexically ill-formed. Because the grammatical feature specification of *haar* has not been properly represented, no A-Chain Condition violation occurs when the pronoun is coindexed with the subject NP. Due to the fact that these two expressions are not coarguments, no Principle B violation occurs either. Note that the hypothesized child LF in (8a) has exactly the same structure as the fully grammatical Dutch LF in (8b), in which *haar* is replaced by the simplex reflexive *zich* 'himself/herself/itself/themselves'.

Philip and Coopmans's (1996) claim, then, is that another kind of child pronoun semantic identity error which is completely independent of Rule I failure can also occur in certain syntactic contexts in certain child languages—and does occur in child Dutch in contexts such as (6b). We call this second hypothesized factor *lexical failure*. Note that lexical failure cannot be a cause of pronoun semantic identity error with sentences such as (6a) or (5a) because Principle B would rule out a bound variable reading of the pronoun even if the pronoun were incorrectly represented as [–R]. Thus, although pronoun semantic identity errors due to Rule I failure may occur with both sentence types in (6), pronoun semantic identity errors due to lexical failure may only occur with the sentence type in (6b). According to Philip and Coopmans, this circumstance is what accounts for the observation that the average frequency of pronoun semantic identity errors with a

group of same-age Dutch children was significantly higher with the latter sentence type than with the former.

There are other contexts in which lexical failure appears to give rise to pronoun semantic identity errors. Prior to Philip and Coopmans's (1996) study, Sigurjónsdóttir and Coopmans (1996) observed that a group of Dutch children showed the pronoun semantic identity errors significantly more often with sentences such as (9a) than with sentences such as (6a) (both presented in embedded contexts). The difference between the two sentence types is that (9a) contains a verb that is lexically ambiguous between a transitive reading and an inherently reflexive reading, whereas the verb in (6a) is unambiguously transitive. Sigurjónsdóttir and Coopmans accounted for their observation as follows: Due to the potential inherent reflexivity in the case of (9a), which causes Principle B to be vacuously satisfied, children who lack knowledge of the [+R] status of pronouns will be able to assign well-formed A-Chains of the sort shown in (9b), in which [+REFL] represents inherent reflexivity. Note that Principle A is also satisfied in (9b), precisely because the predicate is inherently reflexive:

(9) a. Het jongetje wast hem.
'The boy is washing him.'
b. Het jongetje$_1$ wast[+REFL] hem[–R]$_1$.
'The boy is washing himself.'

This gives rise to one sort of pronoun semantic identity error (i.e., the kind caused by lexical failure). On the other hand, children who always fully represent pronouns as [+R] will always counterindex the NPs in (9a), just as they do for sentences such as (6a). However, these children will experience Rule I failure some of the time, giving rise to another sort of pronoun semantic identity error. Thus, just as in the case of the ECM construction of the Philip and Coopmans study, for any group of children of a given age range, a significantly greater number of pronoun semantic identity errors is predicted to occur with sentences like (9a) than with sentences like (6a).

We adopt the basic analysis of Coopmans, Philip, and Sigurjónsdóttir for sentences like (6b) and (9a), and we show that it captures the facts for Norwegian child language as well; however, we make a slightly different claim about what lexical failure consists in. In our view, lexical failure does not occur because children lack knowledge of the adult lexical properties of pronouns, but rather it occurs because, in certain child languages, children often fail to fully integrate their knowledge of the grammatical lexical features of pronouns into their representations of sentences containing pronouns. This hypothesis distinguishes our proposal empirically from that put forth in Sigurjónsdóttir and Coopmans (1996) and Philip and Coopmans (1996) in that our proposal does not predict that an individual child who shows lexical failure on one occasion will necessarily do so on all occasions. In contrast, the claim that lexical failure reflects ignorance of a property of the adult lexicon pre-

TABLE 1
Adult Logical Forms: Antecedent Type × Binding Type

	Local Binding	*Nonlocal Binding*
Quantificational antecedent	*[QP_1 V $pron_1$] (Principle B violation) [QP_1 V $pron_2$] (Rule I irrelevant)	*[QP_1 V [. . . $pron_1$ . . .]] (A-Chain Condition violation) [QP_1 V [. . . $pron_2$ . . .]] (Rule I irrelevant)
Nonquantificational antecedent	*[NP_1 V $pron_1$] (Principle B violation) [NP_1 V $pron_2$] (Rule I blocks coreference)	*[NP_1 V [. . . $pron_1$ . . .]] (A-Chain Condition violation) [NP_1 V [. . . $pron_2$ . . .]] (Rule I blocks coreference)

Note. QP = quantifier phrase; V = verb; NP = noun phrase.

dicts that no individual child should ever show lexical failure on one repeated measure of an experimental condition eliciting it but not on another.

Finally, consider the LFs of the type in (10), which are ungrammatical in every language:

(10) a. *Every boy_1 likes him_1.
b. *Every boy_1 expects him_1 to win.

According to our theoretical assumptions, the only thing that rules out coindexation in (10b) is the A-Chain Condition, whereas (10a) is ruled out both by Principle B and the A-Chain Condition. If well-formed A-Chains overgenerate in child language due to incomplete representation of the grammatrical features of pronouns, then some of the time (10b) should be syntactically grammatical for children as well.[9] Thus, the lexical failure hypothesis makes the novel prediction that pronoun semantic identity errors may also occur with potential antecedents that are quantificational, as long as the potential antecedent and the pronoun are not coarguments. This is the case in which lexical failure and Rule I failure have nonoverlapping effects.

The predictions for child language that our analysis makes thus far can be illustrated by comparison with the predictions it makes for adult linguistic performance. Table 1 summarizes schematically the predictions for adults; Table 2 summarizes the predictions for children. Each cell of the two tables holds two alternative indexing representations for a given string, abstracting away from irrelevant details. Coargument domains are represented by brackets. (Assume also that the sentences have been associated with contexts in which Rule I would not license coreference.)

[9]Under lexical failure, the syntactically well-formed child LF corresponding to (10b) would be [every boy_1 [e_1 expects him[–R]$_1$ to win]], with the A-Chain (e_1,him[–R]$_1$) satisfying the A-Chain Condition. (The A-Chain Condition does not apply to the chain (every boy_1, e_1) because its head occupies an A-bar position.)

TABLE 2
Child Logical Forms Under Lexical Failure (Pronoun = [–R])

	Local Binding	*Nonlocal Binding*
Quantificational antecedent	*[QP_1 V $pron_1$] (Principle B violation) [QP_1 V $pron_2$] (Rule I irrelevant)	[QP_1 V [. . . $pron_1$ . . .]] (A-Chain Condition satisfied) [QP_1 V [. . . $pron_2$ . . .]] (Rule I irrelevant)
Nonquantificational antecedent	*[NP_1 V $pron_1$] (Principle B violation) [NP_1 V $pron_2$] (Rule I failure lets in coreference)	[NP_1 V [. . . $pron_1$ . . .]] (A-Chain Condition satisfied) [NP_1 V [. . . $pron_2$ . . .]] (Rule I failure lets in coreference)

Note. QP = quantifier phrase; V = verb; NP = noun phrase.

In all eight cases in Table 1, the adult grammar either prohibits coreference or limits it to special contexts in which Rule I allows it. For sentences without any kind of licensing context ("out-of-the-blue" sentences), the adult grammar prohibits a coreference interpretation of the matrix subject and the pronoun in all cases.

As can be seen from Table 2, children are predicted to differ from adults in precisely the cases where only the A-Chain Condition plays a crucial role in the adult grammar, in which case lexical failure will determine one type of pronoun semantic identity error. In addition, another type of pronoun semantic identity error will arise due to Rule I failure in the case of LFs containing counterindexed expressions. This section has summarized the way in which our proposal accounts for the findings of prior research. We now turn to Norwegian and show how our analysis makes a new set of predictions concerning the role of lexical failure and the A-Chain Condition, made possible by a special property of Norwegian pronouns.

3. NORWEGIAN PRONOUNS

3.1. Antisubject Orientation

In a language like Dutch, the lexical failure hypothesis can only be tested with verbs that potentially have an inherently reflexive reading or with ECM constructions: two cases in which only the A-Chain Condition determines noncoreference. Norwegian has both of these contexts, but it also has an additional grammatical phenomenon not found in Dutch—namely, *antisubject oriented pronouns* (Hestvik (1992), Vikner (1985)). An antisubject oriented pronoun is a pronoun that cannot have the same semantic value as the closest subject, even though this subject is outside the binding domain of the pronoun. This purely structural condition is illustrated by the sentences in (11), in which binding domains (i.e., coargument domains) are bracketed:

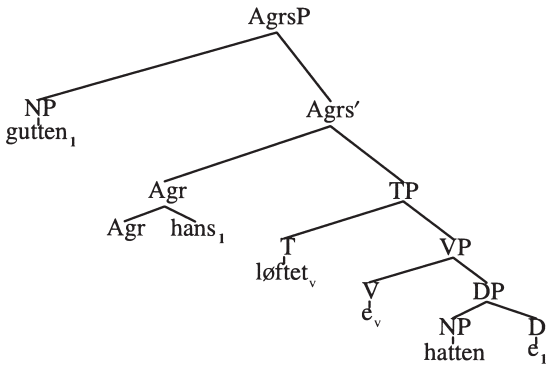
(11) a. *Gutten$_1$ løftet [hatten hans$_1$].
boy-DEF lifted hat his
'The boy lifted his hat.'
b. *Jenten$_1$ satte stolen [bak henne$_1$].
girl-DEF put chair-DEF behind her
'The girl put the chair behind her.'

This effect is not derivable from the theory of Reinhart and Reuland (1993). In all these cases, the pronoun and the ungrammatical antecedent are not coarguments, so Principle B does not apply. Antisubject orientation thus presents a genuine challenge for this theory. Could the A-Chain Condition be responsible for antisubject orientation? Seemingly not, because structurally comparable sentences in English are completely grammatical, as shown in (12):[10]

(12) a. The boy$_1$ lifted [his$_1$ hat].
b. The boy$_1$ looked [behind him$_1$].

We explain this phenomenon as follows: Following Hestvik (1992), we propose that pronouns in Norwegian crucially differ from English in that they undergo obligatory head movement at LF to the functional head of the clause. As a consequence of this movement, the pronoun enters into a local structural relation with the subject of the clause. Thus, according to this theory, the ungrammatical LFs corresponding to (11a) is roughly as shown in (13):

(13)



[10]Conceivably, binding domains in Norwegian could simply be different from those in English by extending beyond the bracketed constituents to the containing clauses. However, the problem for this approach is that a closer nonsubject may bind the pronoun:

(i) Gutten$_1$ viste mannen$_2$ [hatten hans$_{2/*1}$].
boy-DEF showed man-DEF hat-DEF his
'The boy showed the man his hat.'

Our analysis therefore maintains a nonparametrized version of binding domain theory.

The crucial property of (13) is that the subject and the pronoun occur in a local structural relation at LF. Moreover, if subject and pronoun are coindexed, the subject antecedent-governs the pronoun. This gives rise to an interpretive A-Chain headed by the subject NP and tailed by the moved pronoun.[11] Such A-Chains would also be required for representations of Romance reflexive clitic constructions.

The representation in (13) violate the A-Chain Condition if the pronoun is marked [+R]. Thus, the pronoun in fact cannot be coindexed with the subject, and only a representation in which it is counterindexed with the subject will be well-formed. This explains the antisubject-orientation property of the pronouns in the constructions in (11) and in Norwegian in general. In other words, we propose that antisubject orientation is a direct consequence of the A-Chain Condition applying at LF. Due to the fact that this explanation of the phenomenon does not rely on Principle B, the construction no longer poses a problem for the Reinhart and Reuland (1993) theory.

The proposed LF-movement of pronouns is the same as that proposed for reflexives by various authors (e.g., Hestvik (1992), Pica (1987)) to account for the inverse of antisubject orientation, namely the obligatory subject orientation of long-distance reflexives and possessive reflexives. Note that all the ungrammatical cases can be made grammatical by substituting the pronoun with a reflexive, as illustrated in (14):

(14) a. Surface: Gutten$_1$ løftet [hatten sin$_{1/*2}$].
'The boy lifted his hat.'
LF: Gutten$_1$ sin$_1$ [løftet [hatten t$_1$]].
b. Surface: Jenten$_1$ satte stolen [bak seg$_1$].
'The girl put the chair behind her.'
LF: Jenten$_1$ seg$_1$ [satte stolen [bak t$_1$]].

The A-Chains formed at LF in (14a,b) precisely satisfy the A-Chain Condition at LF because the reflexive *seg* and its possessive correlate *sin* are [–R] and can form an interpretive A-Chain with the subject after LF-movement.

Due to the fact that A-Chains are crucially involved in this analysis, our account of antisubject orientation in the adult grammar immediately generates predictions for child language, which we discuss next.

3.2. Predictions for Norwegian Child Language

Suppose Norwegian children, like Dutch children, go through a stage during which pronouns are often misrepresented as [–R]. If this happened with a sen-

[11]We assume that an A-Chain is defined by the syntactic position of its head. Note also that we assume that the relation between the moved element and its trace does not interact with the A-Chain Condition.

tence such as (11a), a syntactically grammatical LF such as (15b) would be assigned to it:

(15) a. Gutten løftet hatten hans.
'The boy lifted his hat.'
b. Gutten$_1$ [hans[–R]$_1$ løftet [hatten t_1]].
c. Gutten$_1$ [hans[–R]$_2$ løftet [hatten t_2]].

The same child capable of assigning the LF (15b) to the string in (15a) some of the time, due to lexical failure, should also be capable of assigning the adult-grammatical LF (15c) to this sentence. When this happens, however, because the counterindexation in (15c) would invoke Rule I, Rule I failure may occur. Thus, for the same child both lexical failure and Rule I failure account for pronoun semantic identity errors with sentences like (15a).[12] The situation is exactly analogous to that of the Dutch ECM case tested by Philip and Coopmans (1996): two independent factors potentially determining pronoun semantic identity errors, so a significantly greater number of such errors are predicted to occur.

On the other hand, if the potential antecedent of the pronoun were quantificational, Rule I would never be invoked, and hence there would be no opportunity for Rule I failure to occur. However, in the case of sentences such as (16a), we still predict that bound variable readings of the pronoun will occur in child Norwegian some of the time as a consequence of lexical failure and assignment of syntactically grammatical (but lexically ill-formed) LFs such as (16b):

(16) a. Hver eneste gutt løftet hatten hans.
'Every boy lifted his hat.'
b. Hver eneste gutt$_1$ [hans[–R]$_1$ løftet [hatten t_1]].

Hence, there would be fewer opportunities for semantic identity errors than with cases like (15), but still a possibility for error via lexical failure.

These predictions are tested in the experiments presented next. In Table 3, we summarize the grammatical status of exemplified construction types in the hypothesized Norwegian child grammatical system, when lexical failure occurs.

[12]Avrutin (1994) hypothesized that only an anaphoric expression interpreted as a bound variable moves at LF, a view that is fully compatible with our proposal. Instead of LF (15c), the following LF would then be generated, with the pronoun *in-situ*:

(i) Gutten$_1$ viste mannen$_2$ [hatten hans$_2$].

In other words, only coindexation with subject NP yields obligatory movement to agreement (Agr). Note that this alternative makes the same predictions, as Rule I failure will occur with (i) just as with (15c). This approach would also solve a problem raised by a reviewer—namely, that if the pronoun moves to Agr in (15c), it would seem to c-command the indirect object, perhaps creating an unwanted A-Chain with that element. This problem does not arise under (i).

TABLE 3
Norwegian Child LFs Under Lexical Failure (Pronoun = [–R])

	Local Binding	*Nonlocal Binding*
Quantificational antecedent	**[Hver eneste gutt$_1$ peker på ham$_1$]*. 'Every boy is pointing at him.' (Principle B violation)	*Hver eneste gutt$_1$ holder [hatten hans$_1$]*. 'Every boy is holding his hat.' *Hver eneste gutt$_1$ satte stolen [bak ham$_1$]*. 'Every boy put the chair behind him.' (A-Chain Condition satisfied)
Nonquantificational antecedent	**[Gutten$_1$ peker på ham$_1$]*. 'The boy is pointing at him.' (Principle B violation) *[Gutten$_1$ peker på ham$_2$]*. (Rule I failure lets in coreference)	*Gutten$_1$ holder [hatten hans$_1$]*. 'The child is holding his hat.' *Gutten$_1$ satte stolen [bak ham$_1$]*. 'The boy put the chair behind him.' (A-Chain Condition satisfied) *Gutten$_1$ holder [hatten hans$_2$]*. *Gutten$_1$ satte stolen [bak ham$_2$]*. (Rule I failure lets in coreference)

The left-hand column illustrates local binding constructions (coargument contexts), and the right-hand column illustrates constructions with nonlocal binding (noncoargument contexts), specifically constructions with possessive pronouns and with pronouns as objects of locative prepositions.

3.3. Principal Experimental Predictions

Table 3 indicates the grammatical status of various LFs in the hypothesized child grammar when lexical failure occurs. How does this translate into experimentally testable predictions about linguistic behavior? In an experimental task in which the child is asked to determine the truth or falsity of a target input sentence of the form [NP V . . . pronoun], we assume that the child's parser will generate for this sentence either a coindexed or a counterindexed LF, the semantic interpretation of which will be the basis for a response in experiment. We also assume that Rule I failure and lexical failure are performance factors whose strength is a function of maturation or linguistic experience. Given these assumptions, our predictions about the relative frequency of nonadultlike responses indicative of a pronoun semantic identity error are shown in Table 4 for each of the principal target input sentence types that are tested.

Specific frequencies are not predicted because we do not know the extent to which lexical failure or Rule I failure can be expected to occur, neither with a particular child nor with a group of children of a given age range. However, with children of a given age range, pronoun semantic identity errors are predicted to occur significantly more often (*a lot*) when both lexical failure and Rule I failure are in principle possible than when only one of these potential factors can possibly affect responses (*some*). In addition, pronoun semantic identity errors are pre-

TABLE 4
Predicted Relative Frequency of Child Pronoun Semantic Identity Errors

	Local Binding	*Nonlocal Binding*
Quantificational antecedent	None	Some
Nonquantificational antecedent	Some	A lot

dicted never to occur when neither of these potential factors can affect responses (*none*).

We now turn to the experiments testing these predictions. Experiment 1 tests the standard predictions of the left column of Table 4 by comparing children's interpretation of pronouns in local contexts when the potential antecedent is referential and when it is quantificational. Experiment 2 examines the extent to which the potential inherent reflexivity of the predicate may affect the frequency of semantic identity errors in local contexts. Experiments 1 and 2 are designed to replicate with Norwegian children the findings of similar experiments carried out in Dutch by Philip and Coopmans (1996) and by Sigurjónsdóttir and Coopmans (1996), and they also replicate the findings that Chien and Wexler (1990) reported for English children. In addition, Experiments 1 and 2 provide a baseline for assessing the extent to which children show pronoun semantic identity errors in antisubject orientation constructions, which are tested in Experiments 3 and 4. Some children who participated in the study participated in all four experiments; however, due to circumstances beyond our control, others did not.[13] Furthermore, although Experiments 2, 3, and 4 were carried out concurrently with one group of children, we present each experiment completely in a separate and easily readable section for expository reasons, and, in section 8, we consider within-subjects comparisons for all experiments.

4. EXPERIMENT 1: LOCAL BINDING

The principal objective of Experiment 1 was to test the basic prediction that in an experimental context Norwegian preschool children should show adultlike comprehension performance significantly more often in local contexts with quantificational potential antecedents than in local contexts with referential potential antecedents. The types of target input sentences used for the two experimental conditions testing this prediction (called QLOCAL and LOCAL) are illustrated in (17) and (18), respectively:

[13]The summer holidays intervened between Experiment 1 and Experiments 2, 3, and 4, and many children who had participated in Experiment 1 did not return in the fall to the preschool that they had been attending the previous spring.

(17) a. Driver hver eneste pike og klapper henne?
do-PRES every single girl and pat-PRES her
'Is every girl patting her?'
b. Peker hver eneste pike på henne?
point-PRES every single girl-DEF at her
'Is every girl pointing at her?'

(18) a. Driver piken og klapper henne?
do-PRES girl-DEF and pat-PRES her
'Is the girl patting her?'
b. Peker piken på henne?
point-PRES girl-DEF at her
'Is the girl pointing at her?'

The periphrastic determiner *hver eneste* 'every' (literally 'every single') in (17) is an unambiguously distributive universal quantifier. The colloquial progressive *driver-og* construction illustrated in (17a) and (18a) was used for some trials of each condition because it yielded more natural-sounding syntax in yes–no questions with inverted verbs, and the simple present was used for the others. Both aspectual forms express approximately the same meaning as English *be-ing*. In both cases a question is formed by moving the verb to the head of a sentence-initial functional projection. For concreteness and finesse, assume the structure of (18a) to be (19):

(19) $[_{CP}$ Driver$_i$ $[_{IP}$ piken $[_{VP}$ e$_i$ og klapper henne]]].

The verb *driver* is semantically inert in this construction and much like English *do* in do–support constructions, and it does not affect the determination of binding domains.

4.1. Participants

Forty-eight monolingual Norwegian children were interviewed for Experiment 1. Data from 4 children were excluded from analysis on the basis of their less-than-perfect performance on 10 different trials of a condition called NOCTRL, which controlled for inattention and failure to master the experimental task.[14] The 44 children who were included in the experiment showed adultlike performance 100% of the time under the NOCTRL Condition. Their ages ranged from 4;5 to 7;4 (M = 6;3). These 44 children can be divided into two chronological age groups: 15 younger children (age range = 4;5–5;11, M = 5;4) and 29 older chil-

[14]Each trial of the NOCTRL Condition elicited an adult "no" response to a question that did not contain a pronoun or a reflexive (e.g., "Is the boy holding an umbrella?" asked of a picture showing a boy who was not holding an umbrella).

dren (age range = 6;0–7;4, *M* = 6;9). Approximately as many boys as girls participated in Experiment 1. The children were drawn from four different preschools located in Bergen, Norway. In addition, 14 adults (*M* = 26;0) participated as a control group.

4.2. Experimental Methodology

The experiment used a version of the Truth–Value Judgment paradigm (for a general assessment of this methodology, see Gordon (1998)). The experiment incorporated key aspects of the design of the fourth experiment of Chien and Wexler (1990) and key aspects of the procedure used in Crain and McKee (1986). Each different trial of an experimental condition consisted of a picture (visual input) matched with a minimal, thematically neutral verbal description of its contents (context-setting input) and a yes–no question about the picture (target input). The visual and context-setting input always established exactly two linguistic antecedents for the pronoun or reflexive in the target input, a *grammatical antecedent* and an *ungrammatical antecedent*. As in Chien and Wexler, the materials were counterbalanced with respect to elicited adult response (yes or no), with respect to picture type (depiction of a reflexive or nonreflexive action) and with respect to the type of anaphoric element in the target input (pronoun or reflexive).

4.2.1. Procedure. The experiment was presented to the child as a "guessing game" in which one experimenter, the "guesser," would make guesses about pictures that he or she could not see. The pictures were manipulated by another experimenter, the "helper," who sat beside the child on one side of a table facing the guesser. The pictures were presented one at a time and were kept out of sight before and after being presented. At each play of the game, the helper held up a picture so that it faced the child and the helper, with the back of it turned toward the guesser. The helper then listed the objects and individuals depicted in the picture, indicating also how many of each type there were, but not indicating which were agents or which were patients. The child was told these were "hints" for the guesser. The helper pointed at each object as its type was mentioned. As the game progressed, these hints were sometimes elicited from the child, the helper pointing at the objects in the picture to prompt their identification by the child. When the contents of the picture had been described in this fashion, the helper gave a final "hint word." This was a verb in infinitival form describing the kind of activity depicted in the picture. At this point, the guesser repeated the context-setting input, as if musing out loud about the contents of the picture while deciding how to guess, and then delivered the target input (i.e., the "guess"). The child's task, then, was to determine whether each guess was correct and respond by saying "yes" or "no." The child also had the secondary task of "rewarding" the guesser for correct guesses by placing a marble in a designated bowl. (At the end of the game, these marbles were counted to determine the guesser's "score.") When the child had

completed both tasks, the helper placed the picture out of sight and proceeded to the next experimental item. Throughout the experiment, the pictures were always hidden from the guesser. The backs of the pictures, where the guesser's lines were written, were always hidden from the child. Both experimenters used normal prosody at all times.[15]

Throughout the experiment, the guesser unobtrusively recorded the child's responses. If a child spontaneously changed his or her response from "yes" to "no," or vice versa, only the last response was recorded. (This was rare.) If a child failed to give any response the first time the target input was presented, the guesser repeated the guess one time. (This occurred extremely rarely, and a single repetition of the target input was always sufficient to elicit a response.) If a spontaneous child explanation of why the guess was wrong revealed that the child had failed to understand the picture as intended, the helper explicitly corrected this misencoding of the visual input, and the guesser repeated the guess.[16]

The adult participants were tested on the experimental materials by a single experimenter, who simply held up the pictures one at a time and read the context-setting and target inputs off the backs. The adults were told that they were being interviewed to test the materials of an experiment for children.

4.2.2. Design. The children were interviewed individually in a quiet area of the preschool where the experiment was conducted. The adults were interviewed collectively in small groups. The children were presented with the experimental materials in two 15-min experimental sessions spaced approximately 2 weeks apart. The materials of each session were counterbalanced in the manner previously described and presented in a single pseudorandom order to all participants (see Appendix). Each experimental session was preceded by an explanation of the experimental task. In addition, in the first experimental session, the experimental task was first practiced with two warm-up items before the experimental materials were presented. The experimenters were advanced undergraduate linguistics students trained in the experimental methodology.

4.2.3. Materials. The experiment consisted of two test conditions, QLOCAL and LOCAL, and six control conditions (24 items total), interspersed

[15]Just as in English, in Norwegian the prosody of a guess question differs from that of an ordinary yes–no question only in that the sentence-final raise is slightly higher in pitch, expressing uncertainty or hesitancy. When presenting the final context-setting input, the guesser used the prosody of reciting a list (i.e., lightly stressing the noun head of each NP and pausing briefly after each NP).

[16]For example, if the target input were *Is the mother pointing at her?*, presented as a guess about a picture showing a mother pointing at a girl, and the child responded negatively and commented "because the mother is pointing at a boy," then the helper would explain that what the child took to be a boy was supposed to be a girl. After this clarification, the guesser would repeat the guess question. (Again, this was extremely rare.)

with 39 filler items[17] and the 10 different trials of the NOCTRL Condition used to screen out inattentive participants. All together, there were 73 experimental items, 48 of which had pronouns or reflexives in their target inputs. For each test and control condition there were 3 different trials with syntactically identical target inputs but with different types of objects and actions depicted in the visual input and referred to by the linguistic input. A unique 21 × 29 cm hand-drawn color picture was used for each trial of each experimental condition. The predicates used for the target inputs of the test and control conditions were *klappe* 'pat', *peke på* 'point at', and *tørke* 'dry off'.[18] The colloquial progressive *driver-og* construction was used with *tørke* and *klappe*; simple present with *peke på* (see Appendix).

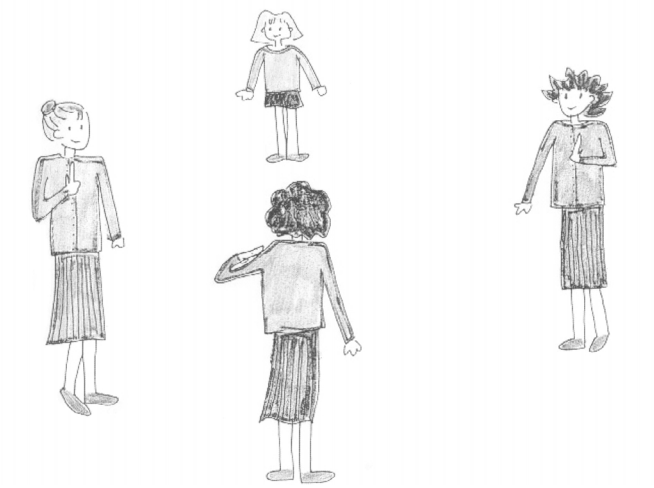
The picture type for the QLOCAL Condition always showed three female individuals of the same age type and one of a different age type (e.g., three girls and one mother; three girls and one grandmother; three mothers and one girl). Trials of the test condition QLOCAL are exemplified in Figure 1. As indicated in parentheses after the condition label in Figure 1, the adult response under the QLOCAL Condition is *Nei* 'No'. All the test conditions of each of the four experiments elicit negative responses from adults.

Trials of the test condition LOCAL are exemplified in Figure 2. For the three different trials of this test condition, the picture showed two female humans of different age types (e.g., a girl and mother; a girl and a grandmother).

Of the six control conditions, two labelled QLOCAL-py and LOCAL-py had structurally identical context-setting and target inputs as the QLOCAL and LOCAL test conditions but were paired with picture types such as those in Figures 3 and 4, so that they elicited adult "yes" responses (e.g., "Is every girl patting her?" asked of the picture in Figure 3 and "Is the girl drying her?" asked of the picture in Figure 4). Two other control conditions labelled QLOCAL-rn and LOCAL-rn are exemplified in Figures 3 and 4. These conditions used the reflexive *seg selv* 'herself/himself/themselves' in the target input rather than a pronoun, and they elic-

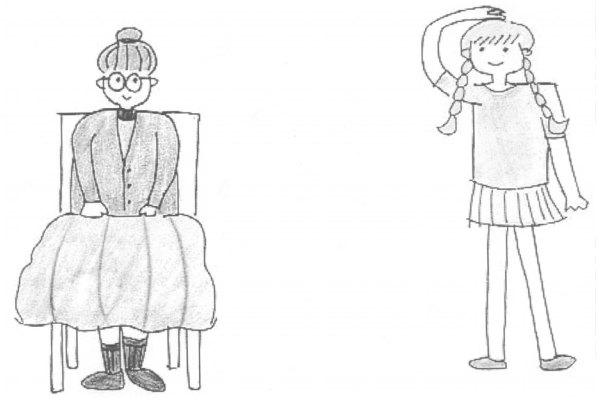
[17]Twelve of the filler items were experimental items for another study (Hestvik and Philip (in press)) examining binding into picture NPs, and another 12 were experimental items a study of binding into perception verb complements. The remaining 15 filler items were target inputs for a study of the "symmetric interpretation" of universal quantifiers (in the sense of Philip (1995)).

[18]The materials for *klappe* and *peke på* were Norwegian versions of materials used in a Dutch experiment reported in Philip and Coopmans (1996). Note that although the predicate *peke på* contains a preposition, this is neither an optional locative preposition nor a separable prefix. As seen by the ungrammaticality of **Jon peker på seg* 'John points at self', *på* of *peke på* does not form a small clause with the NP that follows it, for in this case *seg* would be licensed. (Even in the case of the few verbs in Norwegian which have separable prefixes, the separable prefix does not form a small clause with the following NP, as seen by the ungrammaticality of **Jon pekte ut seg* 'John appointed self'.) *På* of *peke på* is a semantically inert preposition lexically selected by *peke*—just as in English *at* of *look at* is a semantically inert preposition lexically selected by *look*. Therefore, *peke på* is simply a transitive predicate and, as such, is as much subject to Principle B, in the sense of Reinhart and Reuland (1993), as any other transitive predicate. We thank an anonymous reviewer for noting the need for clarification on this point.



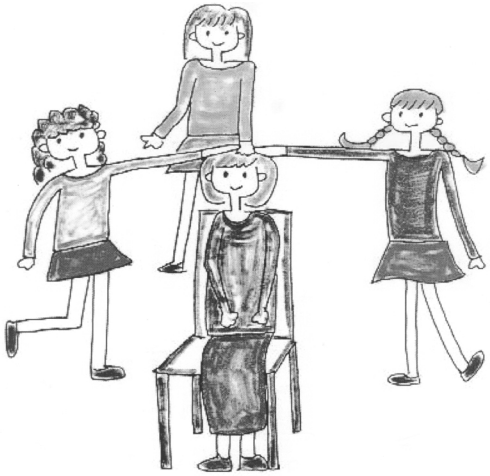
(Context-setting input: Hmm . . . 3 moms and a girl . . .)
Target input: Peker hver eneste mor på henne?
'Is every mom pointing at her?'

FIGURE 1 Example of QLOCAL Condition (Adult answer: No).



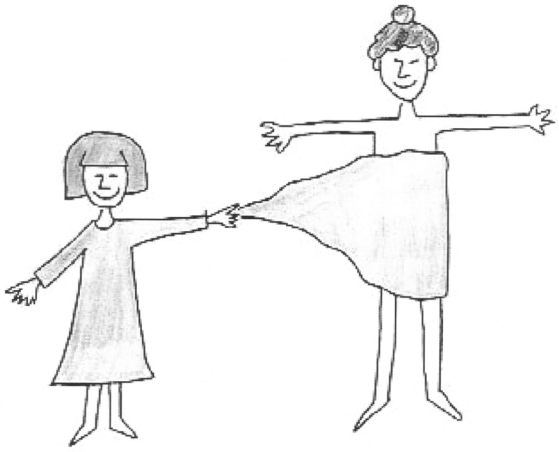
(Context-setting input: Hmmm . . . a girl and a grandma . . .)
Target input: Driver piken og klapper henne?
'Is the girl patting her?'

FIGURE 2 Example of LOCAL Condition (Adult answer: No).



(Context-setting input: Hmm . . . a mom and three girls . . .)
Target input: Driver hver eneste pike og klapper seg selv?
'Is every girl patting herself?'

FIGURE 3 Example of control condition QLOCAL-rn (Adult answer: No).



(Context-setting input: Hmmm . . . a mom and a girl with a towel . . .)
Target input: Driver piken og tørker seg selv?
'Is the girl drying herself off?'

FIGURE 4 Example of control condition LOCAL-rn (Adult answer: No).

ited adult "no" responses. (In general, "-p" means the target input contains a pronoun, "-r" means it contains a reflexive, "-y" means an adult "yes" response is elicited, and "-n" means an adult "no" response is elicited.) The complex reflexive *seg selv* is composed of *selv* 'self' and the simplex anaphor *seg*, which is third person but underspecified for number and gender.

Finally, trials of the QLOCAL-rn and LOCAL-rn Conditions were counterbalanced in turn by the trials of two more control conditions, QLOCAL-ry and LOCAL-ry, which used the same type of context-setting and target inputs as the QLOCAL-rn and LOCAL-rn Conditions but which were paired with pictures of the same type as in Figures 1 and 2 so that they elicited adult "yes" responses (e.g., the question "Is every mom pointing at herself?" asked of the picture in Figure 1).

For all trials of the QLOCAL Condition and its control conditions, the determiner universal quantifier was always *hver eneste* 'every'. For all trials of the experimental conditions testing pronoun comprehension, the third-person feminine pronoun *henne* 'her' was used in the target input. (To provide variety, the target input of some filler items contained a third-person masculine pronoun *ham* 'him'[19].) For the test and control conditions, trials involving adult agents were roughly counterbalanced by trials involving child agents (see Appendix). For all trials of both the test and the control conditions, the context-setting input provided by the guesser always had the form indicated in Figures 1 through 4. Note that no indication is given as to how the two types of objects in the picture are thematically related to each other. Note also that the grammatical antecedent for the pronoun was always mentioned last.

4.3. Results of Experiment 1

The 14 adults of the control group gave the expected "yes" and "no" responses 100% of the time for all test and control conditions. The children's performance was also highly adultlike. In general, the children responded to the target inputs immediately and seemingly without any uncertainty as to the intended meaning or appropriateness of the guesser's guesses.[20] More than 99% of the responses to target inputs containing pronouns or reflexives were immediate "yes" or "no" an-

[19]In colloquial Norwegian, the nominative and the accusative is frequently the syncretic form *han* rather than *ham*. The experimenters sometimes used the syncretic form instead of *ham*.

[20]For a few of the experimental items whose target inputs contained pronouns or reflexives (18 out of a total of 2,112 responses), before responding to the target input the child asked the guesser a clarification question such as *Hvem?* 'Who?', *Henne?* 'Her?', *Seg selv?* 'Herself?', or *Mener du hesten?* 'Do you mean the horse?' In these cases, the target input was simply repeated by the guesser (somewhat infelicitously). In all cases, the child responded "yes" or "no" after a single repetition of the target input. Ten of these responses were adultlike "no" or "yes" responses. Most of the 18 child clarification questions occurred with filler items with target inputs, such as *Is the boy showing the dog a picture of him?* or *Is the boy showing the dog a picture of himself?* A few were also observed with one or another trial of the LOCAL Condition.

TABLE 5
Individual Trial Effects in Experiment 1

Condition	*p Value*	*Condition*	*p Value*
LOCAL-py	.9501	LOCAL	.6205
LOCAL-ry	.9831	LOCAL-rn	.9831
QLOCAL-py	.9831	QLOCAL	.9341
QLOCAL-ry	.9831	QLOCAL-rn	.9831

Note. py = pronoun elicits adult "yes" response; ry = reflexive elicits adult "yes" response; LOCAL = pronoun or reflexive with local nonquantificational antecedent; QLOCAL = pronoun or reflexive with local quantificational antecedent.

swers to the guesser's question.[21] As shown in Table 5 by the p values for Friedman tests comparing performance under each of the three different trials of a given experimental condition, there were no significant effects of individual trials for the test or control conditions. Thus, neither the particular predicates nor the particular pictures used for individual trials significantly affected responses.

As for age effects, the two chronological age groups did not differ significantly in their performance on the test conditions or on the control conditions eliciting negative adult responses. Comparing the performance of the 15 younger children with that of the 29 older children, an analysis of variance (ANOVA) setting the LOCAL, LOCAL-rn, QLOCAL, and QLOCAL-rn Conditions as four levels of the within-subjects factor showed no significant main effect of age ($p \leq .9396$) and no significant Age × Condition interactive effect ($p \leq .8045$). However, an ANOVA taking the four control conditions eliciting adult "yes" responses as levels of the within-subjects factor revealed a significant main effect of age ($p \leq .0339$), but no significant Age × Condition interactive effect ($p \leq .6985$). The main effect here reflects the fact that the older children virtually always gave an adultlike "yes" response under these control conditions, whereas a handful of the younger children gave a nonadultlike "no" response on one or two trials.[22] The average percentages of adultlike "yes" responses for each child chronological age group under each of the control conditions are shown in Table 6. (Standard errors are shown in parentheses in tables with average percentages of responses of a given type.) Due to the fact that there appear to be no age effects with respect to linguistic development, the two child chronological age groups may be treated as a single sample of 44 preschool children. The average percentages of adultlike responses of these 44 children under each control condition are shown in Table 7.

[21]Only very rarely (i.e., for just 24 of the total 3,212 responses) did a child spontaneously change a response from "yes" to "no" (or vice versa) or respond only after a repetition of the target input. For 5 of the total of 3,212 responses, subsequent comments by the child revealed that a "no" response had been determined by a misunderstanding about what the picture was intended to depict. One of these cases involved a filler item; the other 4 involved control items.

[22]This minor age effect would seem merely to reflect a slight difference between the two chronological age groups as regards some secondary, nonlinguistic factor.

TABLE 6
Percent Adultlike "Yes" Responses per Age Group in Experiment 1

Age Group	*LOCAL-py*	*LOCAL-ry*	*QLOCAL-py*	*QLOCAL-ry*
Younger	96% (3)	96% (3)	98% (2)	98% (2)
Older	99% (1)	100%	99% (1)	100%

Note. py = pronoun elicits adult "yes" response; ry = reflexive elicits adult "yes" response; LOCAL = pronoun or reflexive with local nonquantificational antecedent; QLOCAL = pronoun or reflexive with local quantificational antecedent.

TABLE 7
Child Performance on Control Conditions of Experiment 1

Condition	*Target Input*	*% Adultlike Responses*
LOCAL-py (Yes)	Is the girl pointing at her?	98% (1)
LOCAL-rn (No)	Is the girl pointing at herself?	99% (1)
LOCAL-ry (Yes)	Is the girl pointing at herself?	98% (1)
QLOCAL-py (Yes)	Is every girl pointing at her?	98% (1)
QLOCAL-rn (No)	Is every girl pointing at herself?	99% (1)
QLOCAL-ry (Yes)	Is every girl pointing at herself?	99% (1)

Note. py = pronoun elicits adult "yes" response; rn = reflexive elicits adult "no" response; ry = reflexive elicits adult "yes" response; LOCAL = pronoun or reflexive with local nonquantificational antecedent; QLOCAL = pronoun or reflexive with local quantificational antecedent.

TABLE 8
Percent Adultlike Child Responses on Test Conditions of Experiment 1

Age Group	*n*	*Age Range*	*LOCAL*[a]	*QLOCAL*[b]
Younger	15	4;5–5;11	91% (4)	97% (2)
Older	29	6;0–7;4	90% (4)	99% (1)
All	44	4;5–7;4	90% (3)	99% (1)

Note. LOCAL = pronoun or reflexive with local nonquantificational antecedent; QLOCAL = pronoun or reflexive with local quantificational antecedent.

[a]LOCAL (No) = A girl and a mother. . . . Is the girl pointing at her? [b]QLOCAL (No) = Three girls and a mother. . . . Is every girl pointing at her?

(The grammatical adult response is shown in brackets next to the condition label in tables reporting findings for test conditions.) The average percentages of the children's adultlike "no" responses under the two test conditions are shown in Table 8.

Two-tailed sign tests show that the contrast between performance on the LOCAL and QLOCAL Conditions was highly significant for the 44 children taken as a group ($p \leq .0020$) and for the group of 29 older children ($p \leq .0156$). It was not

TABLE 9
Distribution of "No" Responses on LOCAL and QLOCAL Conditions

Condition	*3 No's*	*2 No's*	*1 No*	*0 No's*
LOCAL	33 (75%)	9 (20%)	2 (5%)	
QLOCAL	42 (95%)	2 (5%)		

Note. LOCAL = pronoun or reflexive with local nonquantificational antecedent; QLOCAL = pronoun or reflexive with local quantificational antecedent.

significant for the 15 younger children ($p \leq .2500$). Abstracting away from age effects, for the 44 children as a whole, the distribution of "no" responses under the two test conditions is shown in Table 9. Columns indicate the number of children (percentage of sample in parentheses) showing adultlike performance on three trials, two trials, one trial and no trials, of each test condition.

4.4. Methodological Remarks

Before discussing the findings of Experiment 1, let us consider some of the properties of the experimental methodology that have bearing on the general reliability and soundness of these findings. First, note that the experimental task was straightforward and natural: determine the correctness of a guess about a picture hidden from the guesser. This alone should have reduced to a minimum confounding effects arising from uncontrolled nonlinguistic factors such as inattention, confusion, or misunderstanding about how to complete the experimental task, yes-response bias, and so forth. Second, the children's performance on the 10 NOCTRL items served to identify and screen out any children whose linguistic performance may have been even minimally confounded by such effects.

Third, it is highly unlikely that the linguistic performance of any child was confounded by the effects of infelicity or pragmatic inappropriateness. For presupposition failure to occur under our test conditions, a participant would have to fail to notice that the context provides potential referents for the pronoun in the target input. For such a perceptual failure to occur, the participant would have to ignore both the visual input, in which two types of potential referents for the pronoun are saliently displayed, and the context-setting input, in which two linguistic antecedents for the pronoun are verbally presented. Thus, under standard assumptions about what gives rise to infelicity both for children and adults (Chierchia and McConnell-Ginet (1992)), the target input should have been completely felicitous. Due to the fact that the pictures were always hidden from the guesser, the target input sentences should have been pragmatically appropriate as well, either as guesses or simply as requests for information. (Thanks to the counterbalancing, approximately half the time the guesser made an incorrect guess, so the target input sentences had all the appearance of genuine guesses.) Moreover, even if a

child did somehow come to the conclusion that his or her performance was being examined and that the guesser was only pretending to guess, it is not clear why this would make the target input sentence pragmatically inappropriate; in such a circumstance an adult would find the question pragmatically appropriate as a test question and endeavor to respond as correctly as possible.

Fourth, thanks to the counterbalancing, the children witnessed both three verifying and three falsifying situations for each sentence type that was tested. If any child had general difficulty determining the truth value of a sentence due to a general difficulty imagining counterfactual situations that would satisfy its truth conditions, then the fact that precisely such verifying situation types were presented in trials of the control conditions should have eliminated this hypothetical problem.[23]

Fifth, regarding pronoun comprehension, note that the context-setting input firmly established the linguistic antecedence of both a grammatical and an ungrammatical antecedent. This part of the design removed another possible confounding effect that can arise when pronouns are tested without a grammatical linguistic antecedent.[24] In addition, in the context-setting input the two potential linguistic antecedents for the pronoun were never assigned θ-roles or connected to grammatical functions of a specific predicate. (They were merely listed.) This removed yet another possible confounding effect concerning pronoun comprehension because it made it impossible for the children to respond on the basis of a strategy of parallel thematic role or parallel grammatical function when determining the reference of a pronoun.[25]

In sum, the experimental paradigm was carefully designed to maximally reduce potential confounding effects. Although every psycholinguistic experiment will have some "experimental noise," one would expect the children's linguistic performance in Experiment 1 to be predominantly, if not entirely, determined by their grammatical competence.

4.5. Discussion of Experiment 1

The first important finding is that the children in the study showed highly adultlike performance under all the control conditions. The fact that they responded "no" and

[23]We doubt that any normal preschooler ever experiences such a difficulty. If an ability to imagine counterfactual situations were a cognitive prerequisite for determining the truth value of a sentence, and if preschool children generally lacked this ability, then they would lack the ability to determine whether a sentence was true or false under ordinary conditions of use.

[24]As Grimshaw and Rosen (1990, 200–203) pointed out, in the situation of an experimental context in which the only linguistic antecedent for a pronoun is an ungrammatical antecedent and the participant is forced to choose between pragmatic well-formedness and syntactic well-formedness, it is not inconceivable that participants sometimes choose to violate a syntactic constraint on pronoun reference to achieve pragmatic well-formedness.

[25]Although such strategies are also operant in adult language processing, it has been shown that preschool children rely on them much more heavily than adults for the disambiguation of pronominal discourse anaphora (Karmiloff-Smith (1980)).

"yes" as an adult would under these conditions strongly suggests that (i) their performance was not confounded by the effects of any nonlinguistic factors such as a yes-response bias; (ii) they were attentive and had fully mastered the experimental task; and (iii) they had fully acquired the lexical properties of the verbs and nouns used in the target inputs of trials of the test conditions, and they used these lexical properties exactly as an adult does. Their highly adultlike performance under the LOCAL-rn and LOCAL-ry Conditions shows that they had fully acquired knowledge (and adultlike "obedience") of Principle A of the Binding Theory as well as knowledge of the lexical properties of the Norwegian reflexive *seg selv* that interact with Principle A in single-clause sentences. Their highly adultlike performance under the QLOCAL-rn and QLOCAL-ry Conditions shows that they had no difficulty assigning a bound variable interpretation of a reflexive in a sentence, such as *Every girl is pointing at herself*. This effectively refutes any speculation that a "no" response to a false sentence of the form *Every girl is pointing at her* may have been determined not by the child's knowledge of Principle B, but rather by a general rejection of bound variable readings with determiner universal quantifiers.[26] The children clearly knew that a determiner universal quantifier could have this function when the referentially dependent element was a reflexive. Given the general absence of a yes-response bias, the children's highly adultlike performance under the QLOCAL-py Condition is also consistent with the hypothesis that their rejection of bound variable readings of sentences like *Every girl is pointing at her* was purely due to their complete acquisition (and adultlike obedience) of Principle B of the Binding Theory.

Turning to the findings for the test conditions, the first observation of interest is that the prediction tested by Experiment 1 was borne out: The children showed adultlike performance significantly more often under the QLOCAL Condition than under the LOCAL Condition. Responses indicative of pronoun semantic identity errors were only observed under the LOCAL Condition, in which coreference was possible in principle because the potential linguistic antecedent was a referring expression. Just as predicted by Grodzinsky and Reinhart's (1993) Rule I failure hypothesis, there were no such errors under the QLOCAL Condition in which the potential linguistic antecedent for the pronoun was quantified NP.

Another noteworthy finding of Experiment 1 is the surprising observation that still highly adultlike performance occurred under the LOCAL Condition. Even the group of younger children showed the classic pronoun semantic identity errors with sentences like *The girl is pointing at her* only 9% of the time. This contrasts sharply with the typical performance of, say, English-speaking children, who have repeatedly been shown to exhibit such errors as often as 50% of the time (e.g., Chien and Wexler (1990), Grimshaw and Rosen (1990), McKee (1992)).

[26]This analysis would also face the problem of explaining why the contrast between LOCAL and QLOCAL was statistically more significant with the older children, as if the hypothesized processing problem increased, rather than decreased, with age (cf. Grimshaw and Rosen (1990, 200, 215)).

More telling, the performance of our Norwegian child participants contrasts sharply with the performance of Dutch children of the same age tested on a Dutch version of almost identical materials using an almost identical experimental paradigm (Philip and Coopmans (1996)). This apparent contrast in child languages strongly suggests that the unexpectedly low incidence of pronoun semantic identity errors in simple sentences in Norwegian is due to a grammatical factor (i.e., a basic difference between Norwegian on one hand and Dutch and English on the other).[27]

5. EXPERIMENT 2: REFLEXIVE VERBS

The objective of Experiment 2 was to test the prediction that Norwegian children would show pronoun semantic identity errors significantly more often with sentences such as (20), which contains a verb that is ambiguous between a normal transitive verb and a verb that has an inherently reflexive reading,[28] than with sentences such as (17) and (18), which contain verbs that only have a transitive, nonreflexive reading:

(20) Driver mannen og vasker ham?
do-PRES man-the and wash-PRES him
'Is the man washing him?'

Under an inherently reflexive reading of the verb, a sentence such as (20) would always satisfy both Principles A and B. As discussed, if the inherently reflexive reading of the verb were selected and, thanks to lexical failure, the A-Chain Condition did not block coindexation of the subject and the pronoun, then a pronoun semantic identity error would result. If a transitive reading of the verb were selected instead, then, regardless of whether lexical failure occurred, the NPs would be counterindexed. If Rule I failure then occurred, another type of pronoun semantic identity error some of the time would result. Thus, with a large enough sample, we expect to find a significantly higher average frequency of pronoun semantic identity errors with sentences such as (20) than with sentences such as (18), in which only Rule I failure can determine such errors.

[27]Interestingly, the behavior of the Norwegian children in this experimental condition more closely matches that of Italian children of roughly the same age, who also have been shown to exhibit very low levels of pronoun semantic identity errors in similar experimental contexts (McKee (1992)).

[28]The possibility of a [+refl] reading for the verb *vaske* is demonstrated by the grammaticality and meaning of *Mannen driver og vasker seg* 'The man is washing himself'. This sentence must be interpreted reflexively. *Vaske* also has an ordinary transitive reading, as shown by the grammaticality and meaning of *Mannen driver og vasker ham* 'The man is washing him'. This sentence cannot have a reflexive interpretation.

5.1. Participants

Thirty-four Norwegian children were tested on the materials of Experiment 2. One child was excluded because virtually all responses were "yes." The 33 children who were included in the study were drawn from the same preschools as the children who participated in Experiment 1. Eleven had participated in Experiment 1 4 months prior to participating in Experiment 2. All 33 children simultaneously participated in Experiment 4, and 30 had participated in Experiment 3 2 or 3 weeks prior to participating in Experiment 2. There were approximately as many boys as girls in the study, ranging in age from 4;10 to 6;8 (M = 6;0).[29] There was no adult control group.

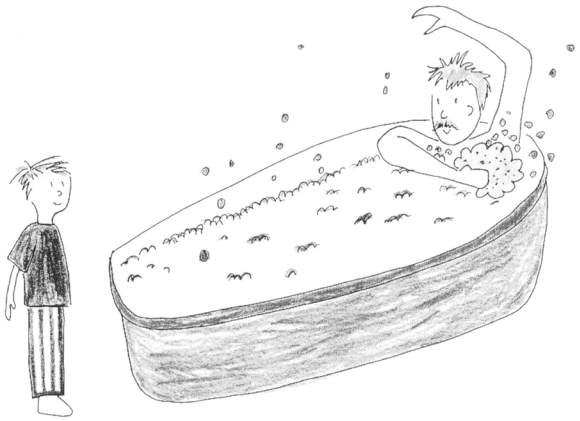
5.2. Experimental Methodology

Experiment 2 used exactly the same experimental methodology as Experiment 1. The design differed slightly in that only 4 NOCTRL items were used to identify and screen out inattentive participants and in that the materials were presented to the children in a single 15-min experimental session. The 12 items of the test and control conditions were interspersed with the 4 NOCTRL items and with the 24 items that constituted the materials of Experiment 4.

5.3. Test and Control Conditions

The test condition REFL is exemplified in Figure 5. The other two trials of the REFL Condition used the predicates *klø* 'scratch' and *sminke* 'put on makeup'. Like *vaske*, each of these predicates has an inherently reflexive reading as well as a transitive reading. The feminine singular accusative pronoun *henne* 'her' was used for two trials of the REFL Condition; the masculine singular accusative pronoun *ham* 'him' was used for one trial. (For more information, see Appendix.) The REFL condition was counterbalanced by three control conditions: REFL-py, REFL-rn, and REFL-ry. As in Experiment 1, REFL-py tested the type of target input used for the REFL condition with pictures that elicited an adult "yes" response, whereas the REFL-rn and REFL-ry conditions tested sentences with reflexives such as *Driver mannen og vasker seg* 'Is the man washing himself' with both falsifying and verifying pictures. For all trials of REFL-rn and REFL-ry, the simplex anaphor *seg* was used.

[29]Five of the children included in the study may have been bilingual to some extent, and the other 28 were known to be monolingual. The performance of these children (children with one or another parent whose first language was not Norwegian) did not differ in any observable respect from that of the monolingual Norwegian-speaking child participants. The other first languages of these possibly bilingual children were Icelandic, German, English, Spanish, and Thai.



(Context-setting input: Hmmm . . . a man and a boy . . .)
Target input: Driver mannen og vasker ham?
'Is the man washing him?'

FIGURE 5 Example of REFL Condition (Adult answer: No).

5.4. Results of Experiment 2

There were no significant effects of neither individual trials nor age.[30] The percentages of adultlike "no" and "yes" responses under the test and control conditions are shown in Table 10. The number of children (percentage of sample in parentheses) showing adultlike performance on three trials, two trials, one trial, and no trials of the REFL Condition are shown in Table 11.

5.5. Discussion of Experiment 2

The principal finding of Experiment 2 is that the prediction regarding potentially inherently reflexive predicates is borne out. The 33 children of Experiment 2 showed pronoun semantic identity errors 32% of the time with sentences like *The man is washing him*. This result contrasts sharply with the performance of the 44 children of Experiment 1, who showed pronoun semantic identity errors only 10% of the time with sentences like *The girl is pointing at her*.

[30]For example, a Friedman test showed the contrasts in performance on the three different trials of the REFL Condition to have a significance of $p \leq .8529$. Regarding age effects, comparing the performance of the 15 younger children ($M = 5;5$) against that of the 18 older children ($M = 6;4$), an ANOVA setting the one test and three control conditions as levels of the within-subjects factor showed the main effect of age group to have a significance of $p \leq .2562$ and the Age × Condition interactive effect to have a significance of $p \leq .6293$.

TABLE 10
Child Performance on Test and Control Conditions of Experiment 2

Condition	*Target Input*	*% Adultlike Responses*
REFL (No)	Is the man washing him?	68% (7)
REFL-py (Yes)	Is the man washing him?	94% (3)
REFL-rn (No)	Is the man washing himself?	94% (3)
REFL-ry (Yes)	Is the man washing himself?	100%

Note. REFL = pronoun or reflexive object of inherently reflexive verb with nonquantificational antecedent; py = pronoun elicits adult "yes" response; rn = reflexive elicits adult "no" response; ry = reflexive elicits adult "yes" response.

TABLE 11
Distribution of "No" Responses on REFL Condition

Condition	*3 No's*	*2 No's*	*1 No*	*0 No's*
REFL	18 (55%)	4 (12%)	5 (15%)	6 (18%)

Note. REFL = pronoun or reflexive object of inherently reflexive verb with nonquantificational antecedent.

Due to the fact that the same experimental methodology was used with both groups of children, and because both groups showed highly adultlike performance under the control conditions associated with each test condition, as well as under the NOCTRL Condition, it seems unlikely that the contrast in performance under the LOCAL and REFL Conditions could be due to nonlinguistic factors. (We return to this issue in section 7.)

The fact that only the feminine pronoun *henne* 'her' was used in trials of the LOCAL Condition, whereas both *henne* and the masculine pronoun *ham* 'him' were used in trials of the REFL Condition, also cannot explain the observed contrast between the LOCAL and REFL Conditions. In the first place, *ham* was used for only one trial of the REFL Condition; for the other two trials *henne* was tested. Second, there were no significant effects of individual trials in Experiment 2. This indicates, among other things, that choice of third-person singular pronoun did not interact significantly with performance on the test conditions of Experiment 2.

Thus, the most plausible explanation of the apparent contrast in performance observed under the LOCAL and REFL Conditions is that lexical failure had an opportunity to occur under the REFL Condition but not under the LOCAL Condition. That is, the descriptive generalization appears to be that Norwegian children are more likely to show pronoun semantic identity errors with predicates that are optionally inherently reflexive than with predicates that are unambiguously transitive and not inherently reflexive (like *peke på* 'point at'), which is exactly what our proposal predicts.

6. EXPERIMENT 3: NONQUANTIFICATIONAL NONLOCAL BINDING

The goal of Experiment 3 was to test the prediction that children would show pronoun semantic identity errors significantly more often with sentences such as in (21), in which the pronoun is not a coargument with the subject, than with sentences such as in (18), in which the pronoun and the subject are coarguments. According to the theory of Reinhart and Reuland (1993), Principle B is (vacuously) satisfied both in the sentences in (21) and in their English analogs. This is why the subject and the pronoun may be semantically identical in English analogs of (21a) and (21b), via binding.[31] In contrast, in adult Norwegian semantic identity of the subject and the pronoun is impossible in (21a) and (21b) because a pronoun coindexed with the subject and moved at LF would result in an ill-formed A-Chain with the subject. Due to the fact that the pronoun is [+R], the A-Chain Condition rules this out, and hence only contraindexing is possible, yielding antisubject orientation:

(21) a. Har mannen lagt kofferten bak ham?
have-PRES man-DEF put-PAST suitcase-DEF behind him
'Has the man put the suitcase behind him?'

b. Løfter mannen hatten hans?
lift-PRES man-DEF hat-DEF his
'Is the man lifting his hat?'

If, at a preschool age, Norwegian children do not consistently represent the lexical properties of pronouns that interact with the A-Chain Condition, then some of the time they should assign nonadultlike, but syntactically grammatical, LFs to sentences such as (21a) and (21b), in which the subject forms an A-Chain with the pronoun after LF-movement. This representation would give rise to a type of pronoun semantic identity error that could never occur with sentences such as in (18), in which Principle B would always block coindexation. In addition, Rule I failure should account for another type of pronoun semantic identity errors with sentences such as (21a) and (21b).

6.1. Participants

Thirty-nine Norwegian children were tested on the materials of Experiment 3. One child was excluded from the study for responding "yes" virtually 100% of the time. The 38 children who were included in the study were drawn from the same

[31]In addition, in the English analogs of (21a) and (21b) the A-Chain Condition is vacuously satisfied as, for structural reasons, no A-Chain can be established between the subject NP and the pronoun in each case.

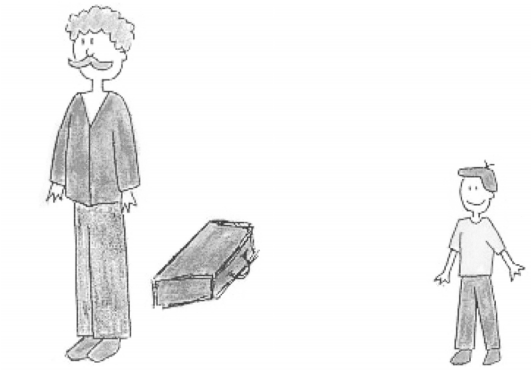
preschools as the children who participated in Experiments 1 and 2. Eleven had participated in Experiment 1 approximately 4 months earlier. There were approximately as many boys as girls in the study, ranging in age from 3;11 to 6;8 (M = 5;10). Five of the children who participated in Experiment 3 may have been to some extent bilingual (see footnote 29). In addition, eight adults (M = 24;0) participated as a control group.

6.2. Experimental Methodology

Experiment 3 used exactly the same experimental methodology as Experiments 1 and 2. The design differed slightly in that 7 NOCTRL items were used to identify and screen out inattentive participants. As in Experiment 2, the materials were presented to the children in a single 15-min experimental session. In addition to the 7 NOCTRL items, there were 24 items for test and control conditions and 12 filler items. The filler items had target inputs such as *Does the boy see him?* or *Does the girl see herself?*

6.3. Test and Control Conditions

The test condition PP, which tested sentences such as (21a), is exemplified in Figure 6. An asterisk in front of the pronoun is used to remind the reader that in adult Norwegian the matrix subject cannot be a grammatical antecedent of the pronoun. (The target input was presented with normal prosody, as in the other experiments.) Two trials of the PP Condition used the predicate *legge* 'put' and the third used the predicate *sette* 'set'. Two trials of the PP Condition tested the masculine



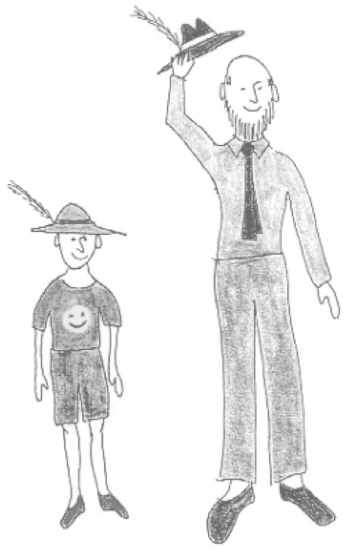
(Context-setting input: Hmmm . . . a man and a boy . . .)
Target input: Har mannen lagt kofferten bak ham?
'Did the man put the suitcase behind *him?'

FIGURE 6 Example of PP Condition (Adult answer: No).

singular accusative pronoun *ham* 'him'; the third trial of this condition tested the feminine singular accusative pronoun *henne* 'her'. (For more information, see Appendix.)

Note that for the PP Condition—and for its three counterbalancing control conditions—the context-setting input provided by the helper had a slightly different form. Instead of simply giving a verb in infinitival form as the "hint word," the helper described an action by the grammatical antecedent and directly asked the guesser to make a specific guess about this action. For example, for the trial of the PP Condition exemplified in Figure 6 the helper said, "The man put the suitcase somewhere. Can you guess where?" This may have introduced a slight bias in favor of adultlike performance. However, it was deemed necessary to control for a possible confounding effect of infelicity. Note that unlike other test conditions, the target input of the PP Condition makes reference to an event that is not fully presented in the visual input because it occurs in the immediate past. The pictures for this condition only showed the outcome or end moment of an event of someone putting something somewhere. If no mention had been made of this past event, the child would have been compelled to infer its existence (accommodation) in order to assign a truth value to the target input.

The test condition POSS, which tested sentences such as (21b), is exemplified in English in Figure 7. Two trials of the POSS Condition used the predicates *se*



(Context-setting input: Hmm . . . a boy and a man . . .)
Target input: Løfter mannen hatten hans?
'Is the man lifting *his hat?'

FIGURE 7 Example of POSS Condition (Adult answer: No).

'see' and *løfte* 'lift', both taking *hatt* 'hat' as object, modified by a possessive pronoun; the third trial used the verb *knytte* 'tie' with *skoene* 'shoes' as object, modified by a possessive pronoun. Two trials of the POSS Condition tested the feminine singular possessive pronoun *hennes* 'her'; the third trial of this condition tested the masculine singular possessive pronoun *hans* 'his'.

Trials of the PP and POSS Conditions were each counterbalanced by trials of three different control conditions, just as in the previous experiments. Trials of the control conditions PP-py and POSS-py used the same type of target inputs as trials of the respective test conditions, matched with pictures eliciting adult "yes" responses (e.g., a picture showing a man putting a suitcase behind a boy; a picture showing a man lifting a boy's hat). Trials of the PP-rn and POSS-rn Conditions used sentences such as in (22a) and (22b), respectively, matched with pictures eliciting adult "no" responses (e.g., a picture showing a man putting a suitcase behind a boy; a picture showing a man lifting a boy's hat). Trials of the PP-ry and POSS-ry Conditions used sentences such in (22a) and (22b), respectively, matched with pictures eliciting adult "yes" responses (e.g., pictures such as in Figures 6 and 7, respectively):

(22) a. Har mannen lagt kofferten bak seg?
have-PRES man-DEF put-PAST suitcase-DEF behind REFL
'Did the man put the suitcase behind himself?'
b. Løfter mannen hatten sin?
lift-PRES man-DEF hat-DEF REFL
'Is the man lifting his own hat?'

6.4. Results of Experiment 3

The eight adults of the control group gave grammatical responses 100% of the time for all the test and control conditions except the POSS-py Condition, in which they gave the expected "yes" response only 96% of the time. (One adult said "no" on one trial of this condition.) As for the children, there were significant effects of neither individual trials nor age group.[32] The percentages of adultlike "no" and "yes" responses under the control conditions are shown in Table 12. The percentages of adultlike "no" responses the children gave under the test conditions PP and POSS are shown in Table 13. The number of children (percentage in

[32]For example, Friedman tests showed the differences in performance on the different trials of the two test conditions, PP and POSS had a significance of $p \leq .3000$ and $p \leq .2884$, respectively. Regarding age effects, comparing the group performance of 18 younger children ($M = 5;4$) against that of 20 older children ($M = 6;4$), an ANOVA taking performance on the PP Condition and its three control conditions as levels of the within-subjects factor showed no significant main effect of age group ($p \leq .2051$) and no significant Age × Condition interactive effect ($p \leq .9630$). A similar ANOVA taking performance on the POSS Condition and its three control conditions as levels of the within-subjects factor also showed no significant main ($p \leq .9997$) or interactive ($p \leq .1595$) age effects.

TABLE 12
Child Performance on Control Conditions of Experiment 3

Condition	*Target Input*	*% Adultlike Responses*
PP-py (Yes)	Did the man put the suitcase behind *him?	61% (6)
PP-rn (No)	Did the man put the suitcase behind himself?	88% (4)
PP-ry (Yes)	Did the man put the suitcase behind himself?	97% (1)
POSS-py (Yes)	Is the man lifting *his hat?	80% (4)
POSS-rn (No)	Is the man lifting his own hat?	86% (4)
POSS-ry (Yes)	Is the man lifting his own hat?	98% (1)

Note. PP = pronoun or reflexive in locative prepositional phrase with nonquantificational antecedent; POSS = possessive pronoun or reflexive with nonquantificational antecedent; py = pronoun elicits adult "yes" response; rn = reflexive elicits adult "no" response; ry = reflexive elicits adult "yes" response.

TABLE 13
Child Performance of Test Conditions of Experiment 3

Condition	*Target Input*	*% Adultlike Responses*
PP (No)	Did the man put the case behind *him?	36% (6)
POSS (No)	Is the man lifting *his hat?	42% (5)

Note. PP = pronoun or reflexive in locative prepositional phrase with nonquantificational antecedent; POSS = possessive pronoun or reflexive with nonquantificational antecedent.

TABLE 14
Distribution of "No" Responses on PP and POSS Conditions

Condition	*3 No's*	*2 No's*	*1 No*	*0 No's*
PP	5 (13%)	9 (24%)	8 (21%)	16 (42%)
POSS	3 (8%)	15 (39%)	9 (24%)	11 (29%)

Note. PP = pronoun or reflexive in locative prepositional phrase with nonquantificational antecedent; POSS = possessive pronoun or reflexive with nonquantificational antecedent.

parentheses) showing adultlike performance on three trials, two trials, one trial, and no trials of the test conditions are shown in Table 14.

6.5. Discussion of Experiment 3

The principal finding of Experiment 3 is that our predictions for Norwegian children's comprehension of the Norwegian analogs of *The man put the suitcase behind him* and *The man is lifting his hat* were robustly borne out. Responses

indicative of a pronoun semantic identity error occurred 64% of the time with the former sentence type and 58% of the time with the latter. If we compare this with the performance of the children of Experiment 1 under the LOCAL Condition, we observe a sharp contrast in the incidence of such errors. Recall that responses indicative of a pronoun semantic identity error were observed only 10% of the time under the LOCAL Condition. Again, given the assumption that Rule I failure is a constant factor across different, same-age groups of children, and given the evidence from performance on the control condition that "experimental noise" due to nonlinguistic factors did not significantly confound the findings of any of our experiments, the observed contrasts between performance under the LOCAL Condition and performance under the PP and POSS Conditions must be attributed to syntactic differences between the sentence types tested under these three test conditions. Specifically, because Principle B is vacuously satisfied, lexical failure has an opportunity to show itself under the PP and POSS conditions. Principle B rules out this possibility in the case of the LOCAL Condition.

In addition, note that pronoun semantic identity errors occurred more often under the PP and POSS Conditions than under the REFL Condition. At first blush, this appears to be slightly problematic for our proposal. We argue that it is not, but for the moment let us postpone discussion of this finding.

7. EXPERIMENT 4: QUANTIFICATIONAL NONLOCAL BINDING

The goal of Experiment 4 was to test our prediction that Norwegian children would show pronoun semantic identity errors significantly less often in contexts such as exemplified in (23), in which the local potential antecedent for the pronoun is a quantificational expression, than in contexts such as exemplified in (21), in which the local potential antecedent was a referring expression. With sentences such as in (23), pronoun semantic identity errors due to Rule I failure should not occur at all because the potential local antecedent is a quantified NP. Thus, the only kind of pronoun semantic identity error that can in principle occur with sentences of the types in (23) are those attributable to lexical failure:[33]

(23) a. Tegner hver eneste dame en sirkel rundt henne?
draw-PRES every single woman a circle around her
'Is every woman drawing a circle around *her?'

[33]We occasionally use a * in front of the pronoun in the English translation to remind the reader that the glossed pronoun is antisubject oriented in adult Norwegian (i.e., ungrammatical if bound by the subject).

b. Holder hver eneste dame i skjørtet hennes?
hold-PRES every single woman in skirt-DEF her-POSS
'Is every woman holding *her skirt?'

7.1. Participants

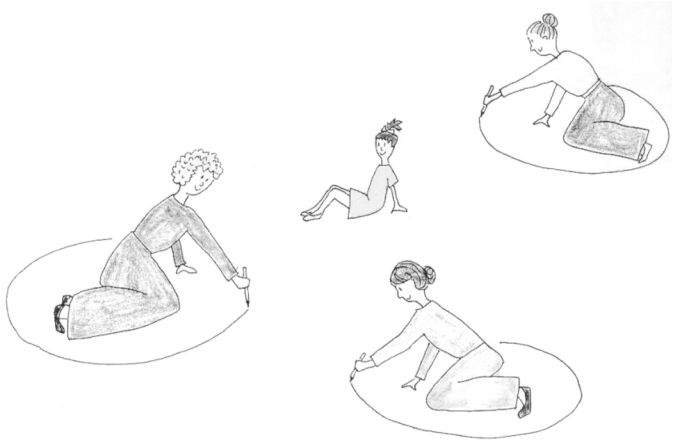
The 33 children who participated in Experiment 4 simultaneously participated in Experiment 2.

7.2. Experimental Methodology

Experiment 4 used exactly the same experimental methodology as the previous experiments (see Experiment 2 and Appendix).

7.3. Test and Control Conditions

The test condition QPP, which tested sentences such as (23a), is illustrated in English in Figure 8. The test condition QPOSS, which tested sentences such as (23b), is illustrated in English in Figure 9. The two other trials of the QPP Condition used the predicates *legge* 'put' and *sette* 'set' and were presented in the special manner discussed in section 6.3 (i.e., with the helper asking for a guess about where objects had been put). The two other trials of the QPOSS Condition used the predicates *se* 'see' and *løfte* 'lift'. Two trials of the QPP Condition tested the pronoun *henne* 'her'; the third trial of this condition tested the pronoun *ham*



(Context-setting input: Hmmm . . . 3 women and a little girl . . .)
Target input: Tegner hver eneste dame en sirkel rundt henne?
'Is every woman drawing a circle around *her?'

FIGURE 8 Example of QPP Condition (Adult answer: No).

(Context-setting input: Hmm . . . 3 ladies and a little girl . . .)
Target input: Holder hver eneste dame i skjørtet hennes?
'Is every lady holding *her skirt?'

FIGURE 9 Example of QPOSS Condition (Adult answer: No).

'him'. Two trials of the QPOSS Condition tested the masculine possessive pronoun *hans* 'his'; the third trial of this condition tested the feminine possessive pronoun *hennes* 'her'.

Trials of the QPP and QPOSS Conditions were each counterbalanced by the trials of three different control conditions, just as in the previous experiments. Trials of the control conditions QPP-py and QPOSS-py used the same type of target inputs as trials of their respective test conditions, matched with pictures eliciting adult "yes" responses (e.g., a picture showing three women drawing concentric circles around one girl; a picture showing three girls each holding on to one woman's skirt). Trials of the QPP-rn and QPOSS-rn Conditions used sentences such in (24a) and (24b), respectively, matched with pictures eliciting adult "no" responses (e.g., a picture showing three women drawing concentric circles around one girl; a picture showing three girls each holding on to one woman's skirt). Trials of the QPP-ry and QPOSS-ry Conditions used sentences such as in (24a) and (24b), respectively, matched with pictures eliciting adult "yes" responses (e.g., pictures such as in Figures 8 and 9, respectively; 'her own' is used as a translation for the reflexive possessive pronoun):

(24) a. Tegner hver eneste dame en sirkel rundt seg?
draw-PRES every single woman a circle around REFL
'Is every woman drawing a circle around herself?'
b. Holder hver eneste jente skjørtet sitt?
hold-PRES every single girl skirt-DEF POSS-REFL
'Is every girl holding her own skirt?'

TABLE 15
Child Performance on Control Conditions of Experiment 4

Condition	*Target Input*	*% Adultlike Responses*
QPP-py (Yes)	Is every woman drawing a circle around *her?	90% (4)
QPP-rn (No)	Is every woman drawing a circle around herself?	88% (5)
QPP-ry (Yes)	Is every woman drawing a circle around herself?	98% (1)
QPOSS-py (Yes)	Is every girl holding *her skirt?	88% (4)
QPOSS-rn (No)	Is every girl holding her own skirt?	93% (3)
QPOSS-ry (Yes)	Is every girl holding her own skirt?	100%

Note. QPP = pronoun or reflexive in locative prepositional phrase with quantificational antecedent; QPOSS = possessive pronoun or possessive reflexive with quantificational antecedent; py = pronoun elicits adult "yes" response; rn = reflexive elicits adult "no" response; ry = reflexive elicits adult "yes" response.

TABLE 16
Percent Adultlike Responses on Test Conditions of Experiment 4

Age Group	*n*	*QPP*[a]	*QPOSS*[b]
Younger	15	60% (9)	53% (11)
Older	18	70% (8)	78% (8)
All	33	66% (6)	67% (7)

Note. QPP = pronoun or reflexive in locative prepositional phrase with quantificational antecedent; QPOSS = possessive pronoun or possessive reflexive with quantificational antecedent.

[a]QPP (No) = Three women and a girl. . . . Is every woman drawing a circle around *her? [b]QPOSS (No) = Three girls and a woman. . . . Is every girl holding *her skirt?

7.4. Results of Experiment 4

There were no significant effects of individual trials.[34] Taking the 15 younger children ($M = 5;5$) as one chronological age group, and the 18 older children ($M = 6;4$) as another, an ANOVA showed that these two age groups differed significantly ($p \leq .0408$) with regard to performance on the QPP Condition and its three associated control conditions. The Age × Condition interactive effect was not significant ($p \leq .5665$), however. Another ANOVA showed a marginally significant main effect of age group ($p \leq .0682$) with respect to performance on the QPOSS Condition and its three associated control conditions. Again, the Age × Condition interactive effect was not significant ($p \leq .1502$). Abstracting away from these main effects of age, the percentages of adultlike "no" and "yes" responses under the control conditions are shown in Table 15. The percentages of adultlike "no" responses that the children gave under the QPP and QPOSS Conditions are shown in Table 16. Abstracting away from age effects, the number of children (percent-

[34]Friedman tests showed that the differences in performance on the different trials of the test conditions QPP and QPOSS had a significance of $p \leq .5666$ and $p \leq .7442$, respectively.

TABLE 17
Distribution of "No" Responses on QPP and QPOSS Conditions

Condition	*3 No's*	*2 No's*	*1 No*	*0 No's*
QPP	14 (43%)	8 (24%)	7 (21%)	4 (12%)
QPOSS	17 (52%)	5 (15%)	5 (15%)	6 (18%)

Note. QPP = pronoun or reflexive in locative prepositional phrase with quantificational antecedent; QPOSS = possessive pronoun or possessive reflexive with quantificational antecedent.

age in parentheses) showing adultlike performance on three trials, two trials, one trial, and no trials of the two test conditions are shown in Table 17.

7.5. Discussion of Experiment 4

The findings of Experiment 4 again robustly bear out the contingent predictions of our analysis for Norwegian children's comprehension of Norwegian sentences comparable to *Every man put a suitcase behind him* and *Every man lifted his hat*. Bound-variable interpretations of the pronoun are unavailable for such sentences in adult Norwegian, yet they appear to be available to Norwegian children approximately one third of the time. Due to the fact that Rule I does not apply at all when a sentence contains only one referential expression (the pronoun), Rule I failure cannot occur in principle under the QPP and QPOSS Conditions. Thus, pronoun semantic identity errors under these two test conditions can only be due to lexical failure. When the child fails to fully represent the grammatical features of the pronoun that interact with the A-Chain Condition, LFs in which the subject binds the pronoun are assigned to target input sentences of the QPP and QPOSS Conditions.

As to the main effect of age that was observed for the QPP and QPOSS Conditions, this would seem to indicate that lexical failure progressively decreases as a function of age (as one would expect).

8. WITHIN-SUBJECTS COMPARISONS

In sections 4 through 7 we presented the findings of four different experiments separately, and we treated the different groups of children that participated in each experiment as validly comparable samples of the same population (i.e., as homogenous groups with respect to all relevant aspects of linguistic and nonlinguistic development). However, as one reviewer pointed out, there is in principle another possible interpretation of our findings—namely, that the cross-experimental contrasts that we observe do not reflect linguistic differences, as we claim, but rather group differences among children of the same age range. On this view, individual differences among children in the same age range can given rise to significant group contrasts. We now consider this issue.

Recall that there were two experimental groups that partially overlapped. The children who participated in Experiment 1 were one of these groups, and those who participated in Experiments 2, 3, and 4 were the other group. Eleven of the children participated in all four experiments. Let us now consider a within-subjects comparison of the performance of these 11 children. Their average percentages of responses indicative of pronoun semantic identity errors under each of the test conditions are shown in Table 18.

For these 11 children, responses indicative of a pronoun semantic identity error occurred significantly more often under the POSS Condition than under the LOCAL Condition (*t* test, $p \leq .0013$; Friedman, $p \leq .0067$). By transitivity, this entails that they also occurred significantly more often under the PP Condition than under the LOCAL Condition. Neither of these findings can be explained in terms of group differences in linguistic development (or in terms of nonlinguistic group differences) because there is only one group of children.[35] In addition, for the 11 children who participated in all four experiments, responses indicative of a pronoun semantic identity error occurred significantly more often under the QPP Condition than under the QLOCAL Condition (*t* test, $p \leq .0107$; Friedman, $p \leq .0348$). Again, by transitivity the contrast between QLOCAL and QPOSS is equally significant. This is another set of false predictions for the view that our findings can be explained in terms of individual differences.

Finally, even the contrast between the LOCAL and REFL Conditions was marginally significant for these 11 children (*t* test, $p \leq .0701$; Friedman, $p \leq .2278$). In sum, none of the predictions of the "group difference" interpretation of our findings are borne out in this within-subjects comparison. Rather, the principal predictions of our analysis are robustly borne out, and none of the predictions of our analysis are falsified by any of the findings for these 11 children.[36] Given the extremely small size of the sample here, the fact that any of our predictions are borne out is remarkable. This attests to the strength of facts that are captured by our analysis. In addition, note that the performance of the 11 children who participated in all four experiments was statistically nondistinct from the performance of the other children in each experimental group of which they were a subset. This is further evidence that the performance of these 11 children is representative of a

[35]Experiment 1, containing the LOCAL Condition, did occur 4 months prior to Experiment 3, containing the PP and POSS Conditions. However, the children showed adultlike performance significantly more often in the earlier experiment. Surely, linguistic development does not consist in a progressive divergence from the target adult grammar.

[36]The contrast between LOCAL and QLOCAL was not significant, but we can deduce from the findings of Experiment 1 that this is merely an artifact of small sample size in the case of the within-subjects comparison. In addition, the contrasts between PP and QPP and between POSS and QPOSS were not significant for the 11 children. However, it becomes clear that these findings are also statistical artifacts when we consider within-subjects comparison of children who participated in Experiments 2, 3, and 4.

TABLE 18
Test Condition Responses Indicative of Pronoun Semantic Identity Error for 11 Children Participating in All Experiments

M Age	*Age Range*	*LOCAL*	*QLOCAL*	*REFL*	*PP*	*POSS*	*QPP*	*QPOSS*
5;11	4;9–6;9	9% (5)	3% (3)	33% (12)	61% (13)	54% (10)	42% (13)	45% (14)

Note. LOCAL = pronoun or reflexive with local nonquantificational antecedent; QLOCAL = pronoun or reflexive with local quantificational antecedent; REFL = pronoun or reflexive object of inherently reflexive verb with nonquantificational antecedent; PP = pronoun or reflexive in locative prepositional phrase with nonquantificational antecedent; POSS = possessive pronoun or reflexive with nonquantificational antecedent; QPP = pronoun or reflexive in locative prepositional phrase with quantificational antecedent; QPOSS = possessive pronoun or possessive reflexive with quantificational antecedent.

TABLE 19
Test Condition Responses Indicative of Pronoun Semantic Identity Error for 30 Children Participating in Experiments 2, 3, and 4

M Age	*Age Range*	*REFL*	*PP*	*POSS*	*QPP*	*QPOSS*
		34% (8)	65% (7)	56% (6)	37% (7)	37% (7)

Note. REFL = pronoun or reflexive object of inherently reflexive verb with nonquantificational antecedent; PP = pronoun or reflexive in locative prepositional phrase with nonquantificational antecedent; POSS = possessive pronoun or reflexive with nonquantificational antecedent; QPP = pronoun or reflexive in locative prepositional phrase with quantificational antecedent; QPOSS = possessive pronoun or possessive reflexive with quantificational antecedent.

single population that is homogenous in all relevant respects (viz., all Norwegian children of the age range we examined).[37]

Putting aside Experiment 1, let us consider a within-subjects comparison of the performance of those children who participated in Experiments 2, 3, and 4. Recall that 30 children did so. The average percentages of responses indicative of pronoun semantic identity errors for these children under the test conditions of Experiments 2, 3, and 4 are shown in Table 19.

Friedman tests show that the contrasts between the PP and QPP Conditions and between the POSS and QPOSS Conditions were both highly significant ($p \leq .0005$ and $p \leq .0176$, respectively). This is just as predicted. The contrast between PP and POSS was not significant (*t* test, $p \leq .1059$), also as predicted. In addition, responses indicative of pronoun semantic identity errors occurred significantly less often under the REFL Condition than under the PP or POSS Conditions (Friedman, $p \leq .0019$ and $p \leq .0176$, respectively). Again, none of these contrasts in performance can be explained by the group difference interpretation of our findings because there is only one group involved. However, they are all exactly as our analysis predicts (except for the REFL vs. PP–POSS contrast).

In conclusion, the results of the two within-subjects comparisons presented in this section empirically refute the speculation that developmental differences between the child groups (or any kind of group differences) may account for the significant contrasts in the frequencies of pronoun semantic identity errors that were observed under the different test conditions of Experiments 1, 2, 3, and 4.

[37]An unpaired *t* test comparing the performance of the 11 children under the REFL Condition against the performance of the other 22 children of Experiment 2 under this test condition shows that the two groups did not differ significantly (pooled, $p \leq .4545$; separate, $p \leq .9197$). The performance of these two groups did not contrast significantly under the QPP Condition (pooled, $p \leq .3681$; separate, $p \leq .4164$) or under the QPOSS Condition (pooled, $p \leq .2209$; separate, $p \leq .2643$). Likewise, unpaired *t* tests show that the performance of the 11 children who participated in all four experiments did not contrast significantly with the performance of the other 27 children of Experiment 3 under the PP Condition (pooled, $p \leq .7192$; separate, $p \leq .7397$) or under the POSS Condition (pooled, $p \leq .6928$; separate, $p \leq .7006$).

9. CONCLUSION

Our predictions have been borne out for all the experimental conditions we tested. For perspicuity, we summarize the findings of the four experiments in Table 20.

The fact that the children showed any pronoun semantic identity errors at all with quantificational antecedents in nonlocal binding contexts (upper right cell of Table 20) is evidence that Rule I failure alone cannot explain the complete distribution of such errors. Rule I does not apply in these contexts, so pronoun semantic identity errors observed cannot be due to a processing failure in the application of Rule I. Our proposal is that the factor giving rise to pronoun semantic identity errors in such contexts is incomplete representation of the grammatical lexical specifications of pronouns that interact with the A-Chain Condition (i.e., lexical failure). When lexical failure occurs, the child is able to construct a bound variable reading of the pronoun, an option not available in the adult grammar for lexical reasons. Note that we assume neither that a given child will either always analyze a pronoun as [+R] or always analyze it as [–R] nor that an abrupt change in knowledge state necessarily occurs when the child consistently represents pronouns as [+R]. Rather, we assume that lexical failure is a performance problem specific to a class of child languages, a problem in retrieving lexical information or in fully integrating it into LFs, and that it may disappear only gradually as a function of linguistic experience (i.e., "practice"), maturational factors, or both. Due to the hypothesis that it is a performance matter, the same child may succeed in constructing a fully adultlike LF some of the time and may fail to do so on other occasions. Recall from Tables 9, 11, 14, and 17 that this is exactly what we found

TABLE 20
Relative Frequency of Pronoun Semantic Identity Errors Predicted and Observed

	Local Binding	*Nonlocal Binding*
Quantificational antecedent	Predicted: None Observed under QLOCAL: 1% (Is every boy pointing at him?)	Predicted: Some Observed under QPP: 34% (Has every girl put a chair behind *her?) Observed under QPOSS: 33% (Has every boy lifted *his hat?)
Nonquantificational antecedent	Predicted: Some Observed under LOCAL: 10% (Is the boy pointing at him?)	Predicted: A lot Observed under PP: 64% (Did the girl put the chair behind *her?) Observed under POSS: 58% (Did the boy lift *his hat?)

Note. QLOCAL = pronoun or reflexive with local quantificational antecedent; QPP = pronoun or reflexive in locative prepositional phrase with quantificational antecedent; QPOSS = possessive pronoun or possessive reflexive with quantificational antecedent; LOCAL = pronoun or reflexive with local nonquantificational antecedent; PP = pronoun or reflexive in locative prepositional phrase with nonquantificational antecedent; POSS = possessive pronoun or reflexive with nonquantificational antecedent.

for some children.[38] Different children may differ in the frequency with which lexical failure generally occurs for them, but with a large enough sample of children of roughly the same age, one should find a uniform average strength of lexical failure for a given age range of children.

In the case of the nonquantificational nonlocal binding context (lower right cell of Table 20), lexical failure can again occur, giving rise to one kind of pronoun semantic identity error; however, when lexical failure happens not to occur, there is still another way in which pronoun semantic identity errors may arise (i.e., via Rule I failure). When the pronoun and the nonlocal antecedent are counterindexed, Rule I will automatically apply. If the child is then unable to complete the procedures accessed when Rule I applies, he or she may allow the pronoun and the NP counterindexed with it to corefer. Thus, because there are two independent factors underlying pronoun semantic identity errors in these contexts, we predict and observe that such errors occur significantly more often in this context than in contexts in which only lexical failure can occur (upper right cell of Table 20) or in contexts in which only Rule I failure can occur (lower left cell of Table 20).

We also correctly predict that when neither Rule I failure nor lexical failure is possible, no pronoun semantic identity errors will be observed (upper left cell of Table 20). The 1% nonadultlike performance observed under the QLOCAL Condition gives us, incidentally, a sense of how minimally our experimental findings are confounded by the "experimental noise" of uncontrolled nonlinguistic factors.

The performance observed under the REFL Condition in Experiment 2, replicating the findings for Dutch of Sigurjónsdóttir and Coopmans (1996), further corroborates our analysis. Recall that in the within-subjects comparison for the 11 children who participated in both Experiment 1 and 2, the contrast between LOCAL and REFL was marginally significant. With a larger sample size, this contrast would no doubt be highly significant, as a glance at the standard errors for these two conditions in Tables 8 and 10 will attest (90% − 6% = 84%, whereas 68% + 14% = 82%). Under our analysis, this contrast is due to the circumstance that lexical failure can only occur under the REFL Condition (because Principle B masks the effect of the A-Chain Condition with normal transitive verbs), whereas Rule I failure can occur under both the LOCAL and the REFL Conditions.

There was, however, one unexpected finding for the REFL Condition. Recall that responses indicative of pronoun semantic identity errors occurred significantly less often under this experimental condition than under the PP or POSS Conditions. At first this seems problematic for our proposal: Why did lexical failure not give rise to pronoun semantic identity errors just as often under all three

[38]Recall as well that a main effect of age was observed in Experiment 4 for the QPP and QPOSS Conditions. This later finding suggests that there is progress, albeit slow, toward adultlike competence in the representation of the lexical features of pronouns that interact with the A-Chain Condition. (The fact that no age effects were observed for the PP and POSS Conditions in Experiment 3 does not falsify this speculative hypothesis because Rule I failure also occurred under these conditions, and we have no independent measure of the rate at which its average frequency diminishes with age.)

test conditions? The answer to this question, we suggest, lies in a more careful consideration of the circumstances that give rise to pronoun semantic identity errors with the three test conditions. Let us assume that indexation is cost free, so that if the processor initially coindexes two NPs and it turns out that this violates some grammatical principle, it can then readily switch to counterindexed representation. Under the PP and POSS Conditions, if the processor initially counterindexed the NPs in the target input, then some of the time an adultlike interpretation would result and, thanks to Rule I failure, some of the time a pronoun semantic identity error would result. If the processor initially coindexed the NPs in the target input, then some of the time the A-Chain Condition would force a switch to counterindexation, with an adultlike interpretation resulting, and sometimes lexical failure would occur, with a pronoun semantic identity error resulting. Crucially, lexical failure is not in any way affected by processing operations; it either freely occurs or does not occur quite independently of grammar.

The situation is not the same in the case of the REFL Condition. Lexical failure interacts with another lexical factor—namely, selection of an inherently reflexive or transitive reading of the verb. Our proposal makes no predictions about the effects of this additional factor,[39] and therefore it is perfectly consistent with the finding that pronoun semantic identity errors occurred less often under the REFL Condition than under the PP and POSS Condition.

Finally, the one major unexpected result of our study was the highly adultlike performance we observed under the LOCAL Condition (i.e., the nonquantificational local binding context; see the lower left cell of Table 20). Although children made significantly more pronoun semantic identity errors in this condition than under the QLOCAL Condition, one would expect Rule I failure to occur more often under the LOCAL Condition than we in fact observed. Aside from noting that this is further evidence of the soundness of our experimental methodology, we have nothing to say about this unexpected observation and offer no explanation of why Norwegian children show such uncharacteristically adultlike performance vis-à-vis, say, Dutch or English children. The descriptive generalization may be stated as follows: Norwegian children hardly ever show pronoun semantic identity errors attributable to Rule I failure in the domain in which Principle B applies, differing in this respect from English and Dutch children and behaving in this respect much like Italian children. It is tempting to attribute this to a parallel in analysis: Norwegian pronouns are like Italian clitics at LF, and semantic properties of clitics prevent them from interacting with Rule I in the same way as full pronouns. We leave this as a speculation.

---

[39]One may, for example, speculate that the lexical ambiguity in the REFL Condition can give rise to a processing strategy that inhibits pronoun semantic identity errors due to lexical failure. Under the REFL Condition, if a child arbitrarily initially selected an inherently reflexive reading of the verb and lexical failure happened not to occur, he or she would have to reject this initial lexical choice, backtrack, and select a transitive reading instead. (Without lexical failure, the LF for an inherently reflexive reading would violate Principle A.) This additional processing may lead some children to adopt a strategy of always selecting the transitive reading first under the REFL Condition.

(Context-setting input: Let's see . . . a lady and a little girl . . .)
Target input: Knytter damen skoene hennes?
'Is the lady tying her shoes?'

FIGURE 10 Example of POSS-py Condition (Adult answer: Yes).

A final piece of support for our overall analysis comes from findings from two control conditions. Recall that performance under the PP-py and POSS-py control conditions of Experiment 3, both of which elicit adult "yes" responses, was significantly nonadultlike in comparison with performance under the other control conditions of this experiment (see Table 12). Adultlike "yes" responses occurred only 61% of the time under the PP-py Condition and only 80% of the time under the POSS-py Condition. An illustration of the POSS-py Condition is given in Figure 10.

Why would a child say "no" rather than "yes" under these control conditions? Clearly, the child's grammar allows assignment of an adultlike LF, which would yield a "yes" response. The answer is that a "no" response is to be expected some of the time under our analysis, on the assumption that lexical failure may also occur occasionally when an adult "yes" response is elicited.[40] There is nothing in the

[40]Apparently, some other factor also affected responses under experimental conditions eliciting an adult affirmative response, such that lexical failure did not occur as often under such experimental conditions as under experimental conditions eliciting an adult "no" response. In Experiment 4 (see Table 15), nonadultlike "no" responses occurred only 10 to 12% of the time under the QPP-py and QPOS-py Conditions, whereas nonadultlike "yes" responses occurred 33% of the time under the QPP and QPOSS Conditions. Moreover, in Experiment 3 (see Tables 12 and 13), negative responses indicative of pronoun semantic identity error occurred only 20% of the time under the POSS-py Condition, but negative responses indicative of this error occurred 58% of the time under the POSS Condition. We have no explanation of this asymmetry and can only speculate that it reflects a grammar-sensitive yes-response bias that may be described as follows: A fully grammatical LF perceived to be true is selected more often than a lexically deviant LF perceived to be false.

experimental context that biases resolution of the pronoun's reference. If lexical failure should occur, the pronoun may be bound by the subject. If so bound, this would yield a "no" response because the truth conditions under such a reading would not be satisfied. For example, in the case of the trial of the POSS-py Condition shown in Figure 10, if lexical failure should occur, that target input sentence would have the same meaning for the child as the Norwegian analog of "Did the lady tie her *own* shoe?"[41]

Anecdotal evidence that this is exactly what happened comes from occasional spontaneous comments by the child as to why the guesser's guess was incorrect under these control conditions. One such production is given in (25):

(25) Child: Nei . . . skoen til jenten!
'No . . . (it's) the girl's shoe!'

The answer indicates that the representation underlying the child's negative response is the unadultlike (26b), whereas the adult using representation (26a) would be compelled to answer in the affirmative:

(26) a. $[_{CP}$ knytter$_v$ $[_{IP}$ damen$_1$ $[_{Infl}$, Infl+hennes$_2$[+R] $[_{VP}$ e$_v$ skoene t$_2$]]]].
b. $[_{CP}$ knytter$_v$ $[_{IP}$ damen$_1$ $[_{Infl}$, Infl+hennes$_1$[–R] $[_{VP}$ e$_v$ skoene t$_1$]]]].

In either case, the response is in fact correct, from the point of view of semantic interpretation of the respective LF representations. This could not be explained by any other theory we know of.

In conclusion, our experimental findings provide strong psycholinguistic evidence in support of the binding theory of Reinhart and Reuland (1993) augmented by Hestvik's (1992) proposal regarding the LF movement of Norwegian pronouns. This theory accounts for adult Norwegian linguistic performance (intuitions) and for nonadultlike linguistic performance in child Norwegian, when plausible performance-theoretical and lexical-developmental assumptions are made about the material causes of the child comprehension error in question. In addition, our lexical failure hypothesis opens a new area of crosslinguistic research on the development of pronominal anaphora. Does lexical failure occur in child English as well? Would it be found, for example, with verbs like *wash* and in ECM constructions? Finally, and perhaps most important, because our analysis of the so-called delay of Principle B in child Norwegian resolves an apparent discontinuity between child and adult grammar into two performance effects, Rule I fail-

[41]That is, (i), with a focus particle as in the English example:

(i) Knyttet damen sin égen sko?
tied the lady her-REFL own shoe
'Did the lady tie her OWN shoe?'

ure and lexical failure, it provides an additional argument in support of the Strong Continuity Hypothesis of first-language acquisition theory.

ACKNOWLEDGMENTS

This research was financed by an academic exchange agreement between University of Bergen and Utrecht University; by a grant from the Faculty of Arts, University of Bergen; and by a visiting professorship to William Philip at the University of Bergen, all of which we gratefully acknowledge. Parts of this work have been previously presented at the 1996 Boston University Conference on Language Development in Boston, the 27th North East Linguistic Society meeting in Montreal, and the 1996 Generative Approaches to Language Acquisition Conference in Edinburgh.

We thank the children who participated in the experiments, their parents, and their teachers at the preschools and kindergartens Slettebakken, Øvrebø, and Sletten, and Slettebakken Skole. We also thank our team of research assistants. Finally, we thank the anonymous reviewers for their helpful comments. Both authors contributed equally to this work.

REFERENCES

Avrutin, S. (1994) "The Structural Position of Bound Variables in Russian," *Linguistic Inquiry* 25, 709–728.

Berman, S. and A. Hestvik (1997) "Split Antecedents, Noncoreference, and DRT," in H. Bennis, P. Pica, and R. Rooryck, eds., *Atomism and Binding*, Foris, Dordrecht, The Netherlands.

Borer, H. and K. Wexler (1987) "The Maturation of Syntax," in T. Roeper and E. Williams, eds., *Parameter Setting*, Reidel, Dordrecht, The Netherlands.

Chien, Y-C. and K. Wexler (1990) "Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics," *Language Acquisition* 1, 225–295.

Chierchia, G. and S. McConnell-Ginet (1992) *Meaning and Grammar: An Introduction to Semantics*, MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1986) *Barriers*, MIT Press, Cambridge, Massachusetts.

Crain, S. and C. McKee (1986) "Acquisition of Structural Restrictions on Anaphora," in S. Berman, J.-W. Choe, and J. McDonough, eds., *Proceedings of the North Eastern Linguistic Society 16*, GLSA, University of Massachusetts, Amherst.

Gordon, P. (1998) "The Truth-Value Judgment Task," in D. McDaniel, C. McKee, and H. S. Cairns, eds., *Methods of Assessing Children's Syntax*, MIT Press, Cambridge, Massachusetts.

Grimshaw, J. and S. Rosen (1990) "Knowledge and Obedience: The Developmental Status of the Binding Theory," *Linguistic Inquiry* 21, 187–222.

Grodzinsky, Y. and T. Reinhart (1993) "The Innateness of Binding and Coreference: A Reply to Grimshaw and Rosen," *Linguistic Inquiry* 24, 69–101.

Grodzinsky, Y., K. Wexler, Y-C. Chien, S. Marakovitz, and J. Solomon (1993) "The Breakdown of Binding Relations," *Brain and Language* 45, 396–422.

Heim, I. (1993) "Anaphora and Semantic Interpretation: A Reinterpretation of Reinhart's Approach," SfS–Report–07–93, University of Tübingen, Germany.

Hestvik, A. (1992) "LF Movement of Pronouns and Antisubject Orientation," *Linguistic Inquiry* 23, 557–594.

Hestvik, A. and W. Philip (in press) "Experimental Evidence for the Difference Between Syntactic and Logophoric Reflexive Binding," in J. Huang, P. Cole, and G. Hermon, eds., *Long Distance Reflexives: Syntax and Semantics*, Academic Press, San Diego.

Karmiloff-Smith, A. (1980) "Psychological Processes Underlying Pronominalisation and Non-pronominalisation in Children's Connected Discourse," in J. Kreiman and A. E. Ojeda, eds., *Papers from the Parasession on Pronouns and Anaphora*, Chicago Linguistic Society, University of Chicago, Illinois.

McKee, C. (1992) "A Comparison of Pronouns and Anaphors in Italian and English Acquisition," *Language* 2, 21–54.

Philip, W. (1995) *Event Quantification in the Acquisition of Universal Quantification*, Doctoral dissertation, University of Massachusetts, Amherst.

Philip, W. and P. Coopmans (1996) "The Role of Referentiality in the Acquisition of Pronominal Anaphora," in K. Kusumoto, ed., *Proceedings of the North East Linguistic Society 26*, GLSA, University of Massachusetts, Amherst.

Pica, P. (1987) "On the Nature of the Reflexivization Cycle," in *Proceedings of NELS 17*, GLSA, University of Massachusetts, Amherst.

Reinhart, T. (1983) "Coreference and Bound Anaphora: A Restatement of the Anaphora Questions," *Linguistics and Philosophy* 6, 47–88.

Reinhart, T. and E. Reuland (1993) "Reflexivity," *Linguistic Inquiry* 24, 657–720.

Rosen, S. and J. Rosen (1995) "Binding, Coreference and Evidence: A Reply to Grodzinsky and Reinhart," paper presented at 1994 CLRF, Stanford University, Stanford.

Safir, K. (1997) "Symmetry and Unity in the Theory of Anaphora," in H. Bennis, P. Pica, and R. Rooryck, eds., *Atomism and Binding*, Foris, Dordrecht, The Netherlands.

Sigurjónsdóttir, S. and P. Coopmans (1996) "The Acquisition of Anaphoric Relations in Dutch," *Amsterdam Series on Child Language Development* 5.

Vikner, S. (1985) "Parameters of Binder and of Binding Category in Danish," *Working Papers in Scandinavian Syntax* 23.

Wexler, K. and Y.-C. Chien (1985) "The Development of Lexical Anaphors and Pronouns," *Papers and Reports on Child Language Development* 24, 138–149.

APPENDIX

Materials for Experiment 1

Session I (Adult Response in Parentheses)

1. FILLER (Yes)	Holder hver eneste elefant en ballong? 'Is every elephant holding a balloon?'
2. NOCTRL (No)	Holder gutten en paraply? 'Is the boy holding an umbrella?'
3. LOCAL-py (Yes)	Driver bestemoren og tørker henne? 'Is the grandmother drying her?'
4. FILLER (No)	Drikker hver eneste gutt en brus? 'Is every boy drinking a soda?'
5. FILLER (No)	Holder en gutt hver eneste paraply? 'Is a boy holding every umbrella?'

6. FILLER (Yes)	Driver politimannen og viser gutten et bilde av seg? 'Is the policeman showing the boy a picture of himself?'
7. NOCTRL (No)	Holder gutten en paraply? 'Is the boy holding an umbrella?'
8. LOCAL-ry (Yes)	Peker mammaen på seg selv? 'Is the mother pointing at herself?'
9. FILLER (Yes)	Driver hunden og viser gutten et bilde av ham? 'Is the dog showing the boy a picture of him?'
10. FILLER (Yes)	Har hver eneste gutt løftet opp en gris? 'Has every boy lifted a pig?'
11. LOCAL-rn (No)	Driver piken og klapper seg selv? 'Is the girl patting herself?'
12. FILLER (Yes)	Holder en elefant hver eneste ballong? 'Is an elephant holding every balloon?'
13. QLOCAL-py (Yes)	Driver hver eneste mamma og tørker henne? 'Is every mother drying her?'
14. FILLER (No)	Holder hver eneste gutt en paraply? 'Is every boy holding an umbrella?'
15. LOCAL (No)	Peker bestemoren på henne? 'Is the grandmother pointing at her?'
16. FILLER (Yes)	Holder hver eneste gris et eple? 'Is every pig holding an apple?'
17. FILLER (Yes)	Holder en giraff hver eneste ballong? 'Is a giraffe holding every balloon?'
18. QLOCAL-rn (No)	Driver hver eneste pike og klapper seg selv? 'Is every girl patting herself?'
19. FILLER (Yes)	Ser moren henne danse? 'Does the mother see her dance?'
20. FILLER (Yes)	Ser piken seg selv danse? 'Does the girl see herself dance?'
21. FILLER (Yes)	Driver hver eneste gutt og rir på en elefant? 'Is every boy riding an elephant?'
22. FILLER (No)	Driver gutten og viser hesten et bilde av ham? 'Is the boy showing the horse a picture of him?'
23. NOCTRL (No)	Holder hesten en ballong? 'Is the horse holding a balloon?'
24. FILLER (No)	Holder en elefant hver eneste ballong? 'Is an elephant holding every balloon?'
25. FILLER (Yes)	Holder hver eneste giraff en ballong? 'Is every giraffe holding a balloon?'
26. FILLER (Yes)	Holder en gutt hver eneste paraply? 'Is a boy holding every umbrella?'
27. FILLER (No)	Ser piken seg selv blåse såpebobler? 'Does the girl see herself blow soap bubbles?'
28. FILLER (No)	Holder hver eneste dinosaur en ballong? 'Is every dinosaur holding a balloon?'

29. FILLER (No)	Ser piken henne hoppe tau? 'Does the girl see her jump rope?'
30. QLOCAL (No)	Driver hver eneste mamma og tørker henne? 'Is every mother drying her?'
31. QLOCAL-ry (Yes)	Driver hver eneste pike og klapper seg selv? 'Is every girl patting herself?'
32. FILLER (Yes)	Driver hver eneste gutt og rir på en hest? 'Is every boy riding a horse?'
33. FILLER (No)	Driver hunden og viser gutten et bilde av seg? 'Is the dog showing the boy a picture of himself?'
34. NOCTRL (No)	Holder gutten en gris? 'Is the boy holding a pig?'
35. FILLER (No)	Holder en giraff hver eneste ballong? 'Is a giraffe holding every balloon?'

Session II

36. FILLER (Yes)	Ser piken henne hoppe tau? 'Does the girl see her jump rope?'
37. FILLER (Yes)	Ser moren seg selv hoppe tau? 'Does the mother see herself jump rope?'
38. QLOCAL-rn (No)	Driver hver eneste pike og tørker seg selv? 'Is every girl drying herself?'
39. LOCAL-py (Yes)	Peker mammaen på henne? 'Is the mother pointing at her?'
40. NOCTRL (No)	Holder piken en paraply? 'Is the girl holding an umbrella?'
41. FILLER (No)	Driver hunden og viser gutten et bilde av ham? 'Is the dog showing the boy a picture of him?'
42. LOCAL-rn (No)	Driver piken og tørker seg selv? 'Is the girl drying herself?'
43. FILLER (No)	Driver politimannen og viser gutten et bilde av seg? 'Is the policeman showing the boy a picture of himself?'
44. LOCAL-ry (Yes)	Driver bestemoren og klapper seg selv? 'Is the grandmother patting herself?'
45. QLOCAL-py (Yes)	Peker hver eneste pike på henne? 'Is every girl pointing at her?'
46. FILLER (No)	Ser piken seg selv hoppe tau? 'Does the girl see herself jump rope?'
47. FILLER (Yes)	Driver gutten og viser hesten et bilde av ham? 'Is the boy showing the horse a picture of him?'
48. NOCTRL (No)	Drikker hesten en brus? 'Is the horse drinking a soda?'

49. LOCAL (No)	Driver piken og klapper henne? 'Is the girl patting her?'
50. QLOCAL-ry (Yes)	Driver hver eneste pike og tørker seg selv? 'Is every girl drying herself?'
51. NOCTRL (No)	Driver mammaen og drikker en brus? 'Is the mother drinking a soda?'
52. FILLER (No)	Ser mammaen henne danse? 'Does the mother see her dance?'
53. FILLER (Yes)	Driver gutten og viser hesten et bilde av seg? 'Is the boy showing the horse a picture of himself?'
54. QLOCAL (No)	Driver hver eneste pike og klapper henne? 'Is every girl patting her?'
55. FILLER (Yes)	Ser piken henne blåse såpebobler? 'Does the girl see her blow soap bubbles?'
56. FILLER (Yes)	Ser bestemoren seg selv blåse såpebobler?' 'Does the grandmother see herself blow soap bubbles?'
57. QLOCAL-rn (No)	Peker hver eneste mor på seg selv? 'Is every mother pointing at herself?'
58. LOCAL-py (Yes)	Driver piken og klapper henne? 'Is the girl patting her?'
59. NOCTRL (No)	Driver bestemoren og drikker en brus? 'Is the grandmother drinking a soda?'
60. FILLER (No)	Driver politimannen og viser gutten et bilde av ham? 'Is the policeman showing the boy a picture of him?'
61. LOCAL-rn (No)	Peker piken på seg selv? 'Is the girl pointing at herself?'
62. FILLER (No)	Driver gutten og viser hesten et bilde av seg? 'Is the boy showing the horse a picture of himself?'
63. LOCAL-ry (Yes)	Driver bestemoren og tørker seg selv? 'Is the grandmother drying herself?'
64. QLOCAL-py (Yes)	Driver hver eneste pike og klapper henne? 'Is every girl patting her?'
65. FILLER (No)	Ser mammaen seg selv danse? 'Does the mother see herself dance?'
66. FILLER (Yes)	Driver politimannen og viser gutten et bilde av ham? 'Is the policeman showing the boy a picture of him?'
67. NOCTRL (No)	Drikker gutten en brus? 'Is the boy drinking a soda?'
68. LOCAL (No)	Driver piken og tørker henne? 'Is the girl drying her?'
69. QLOCAL-ry (Yes)	Peker hver eneste pike på seg selv? 'Is every girl pointing at herself?'
70. NOCTRL (No)	Holder bestemoren en ballong? 'Is the grandmother holding a balloon?'

71. FILLER (No) Ser piken henne blåse såpebobler?
'Does the girl see her blow soap bubbles?'
72. FILLER (Yes) Driver hunden og viser gutten et bilde av seg?
'Is the dog showing the boy a picture of himself?'
73. QLOCAL (No) Peker hver eneste mamma på henne?
'Is every mother pointing at her?'

Materials for Experiments 2 and 4

1. REFL-ry (Yes) Driver jenten og vasker seg?
'Is the girl washing herself?'
2. NOCTRL (No) Holder gutten en paraply?
'Is the boy holding an umbrella?'
3. QPOSS-ry (Yes) Holder hver eneste gutt i buksen sin?
'Is every boy holding his pants?'
4. REFL-rn (No) Driver mannen og vasker seg?
'Is the man washing himself?'
5. QPOSS-py (Yes) Løfter hver eneste mann hatten hans?
'Is every man lifting his hat?'
6. QPP (No) Tegner hver eneste dame en sirkel rundt henne?
'Is every woman drawing a circle around her?'
7. REFL-rn (No) Driver moren og klør seg?
'Is the mother scratching herself?'
8. QPP-ry (Yes) Har hver eneste dame lagt en hårbørste bak seg?
'Has every woman put a hair brush behind her?'
9. NOCTRL (No) Holder hesten en ballong?
'Is the horse holding a balloon?'
10. REFL-py (Yes) Driver mannen og klør ham?
'Is the man scratching him?'
11. QPOSS (No) Ser hver eneste mann på ballongen hans?
'Is every man looking at his umbrella?'
12. QPP-py (Yes) Tegner hver eneste jente en sirkel rundt henne?
'Is every girl drawing a circle around her?'
13. QPOSS-rn (No) Holder hver eneste gutt i buksen sin?
'Is every boy holding his pants?'
14. REFL (No) Driver jenten og sminker henne?
'Is the girl putting makeup on her?'
15. QPOSS-ry (Yes) Ser hver eneste jente på paraplyen sin?
'Is every girl looking at her umbrella?'
16. QPP-rn (No) Tegner hver eneste gutt en sirkel rundt seg?
'Is every boy drawing a circle around himself?'
17. REFL-py (Yes) Driver moren og sminker henne?
'Is the mother putting makeup on her?'

18. QPP-rn (No)	Har hver eneste mann satt en stol bak seg? 'Has every man put a chair behind himself?'
19. QPOSS-py (Yes)	Holder hver eneste jente i skjørtet hennes? 'Is every girl holding her skirt?'
20. NOCTRL (No)	Holder piken en paraply? 'Is the girl holding an umbrella?'
21. QPP-py (Yes)	Har hver eneste gutt lagt en bamse bak ham? 'Has every boy put a teddy bear behind him?'
22. QPOSS-rn (No)	Ser hver eneste dame på paraplyen sin? 'Is every lady looking at her umbrella?'
23. REFL-ry (Yes)	Driver moren og sminker seg? 'Is the mother putting makeup on herself?'
24. QPOSS (No)	Holder hver eneste dame i skjørtet hennes? 'Is every lady holding her skirt?'
25. QPP-ry (Yes)	Tegner hver eneste gutt en sirkel rundt seg? 'Is every boy drawing a circle around himself?'
26. REFL (No)	Driver mannen og vasker ham? 'Is the man washing him?'
27. QPOSS-py (Yes)	Ser hver eneste gutt på ballongen hans? 'Is every boy looking at his balloon?'
28. REFL-rn (No)	Driver jenten og sminker seg? 'Is the girl putting makeup on herself?'
29. QPP (No)	Har hver eneste mann lagt en bamse bak ham? 'Has every man put a teddy bear behind him?'
30. QPOSS-ry (Yes)	Løfter hver eneste jente hatten sin? 'Is every girl lifting her hat?'
31. NOCTRL (No)	Drikker hesten en brus? 'Is the horse drinking a soda?'
32. QPP (No)	Har hver eneste jente satt en stol bak henne? 'Has every girl put a chair behind her?'
33. REFL-ry (Yes)	Driver gutten og klør seg? 'Is the boy scratching himself?'
34. QPOSS (No)	Løfter hver eneste gutt hatten hans? 'Is every boy lifting his hat?'
35. QPOSS-rn (No)	Løfter hver eneste dame hatten sin? 'Is every lady lifting her hat?'
36. REFL (No)	Driver jenten og klør henne? 'Is the girl scratching her?'
37. QPP-ry (Yes)	Har hver eneste gutt satt en boks bak seg? 'Has every boy put a box behind himself?'
38. QPP-py (Yes)	Har hver eneste dame satt en boks bak henne? 'Has every lady put a box behind her?'
39. QPP-rn (No)	Har hver eneste jente lagt en hårbørste bak seg? 'Has every girl put a hair brush behind herself?'
40. REFL-py (Yes)	Driver moren og vasker henne? 'Is the mother washing her?'

Materials for Experiment 3

1. POSS-py (Yes)	Løfter mannen hatten hans? 'Is the man lifting his hat?'
2. FILLER (Yes)	Ser damen seg selv? 'Does the woman see herself?'
3. NOCTRL (No)	Holder gutten en paraply? 'Is the boy holding an umbrella?'
4. PP (No)	Har gutten satt stolen bak ham? 'Has the boy put the chair behind him?'
5. POSS-ry (Yes)	Knytter damen skoene sine? 'Is the lady tying her shoes?'
6. FILLER (No)	Ser mannen seg selv? 'Does the man see himself?'
7. POSS (No)	Ser jenten på hatten hennes? 'Is the girl looking at her hat?'
8. PP-ry (Yes)	Har gutten satt vannkannen bak seg? 'Has the boy put the watering can behind himself?'
9. FILLER-py (Yes)	Ser damen henne? 'Does the girl see her?'
10. PP-py (Yes)	Har jenten satt bøtten bak henne? 'Has the girl put the bucket behind her?'
11. NOCTRL (No)	Holder hesten en ballong? 'Is the horse holding a balloon?'
12. POSS-rn (No)	Løfter mannen hatten sin? 'Is the man lifting his hat?'
13. FILLER (No)	Ser mannen ham? 'Does the man see him?'
14. PP-rn (No)	Har bestefaren lagt ballen bak seg? 'Has the grandfather laid the ball behind himself?'
15. POSS-ry (Yes)	Ser jenten på hatten sin? 'Is the girl looking at her hat?'
16. FILLER (Yes)	Ser mannen seg selv? 'Does the man see himself?'
17. NOCTRL (No)	Holder piken en paraply? 'Is the girl holding an umbrella?'
18. PP-rn (No)	Har jenten satt boksen bak seg? 'Has the girl put the box behind herself?'
19. POSS-py (Yes)	Knytter damen skoene hennes? 'Is the lady tying her shoes?'
20. FILLER (No)	Ser jenten seg selv? 'Does the girl see herself?'
21. PP-py (Yes)	Har bestefaren lagt ballen bak ham? 'Has the grandfather put the ball behind him?'

22. NOCTRL (No)	Drikker hesten en brus?
	'Is the horse drinking a soda?'
23. POSS-rn (No)	Ser jenten på hatten sin?
	'Is the girl looking at her hat?'
24. FILLER-py (Yes)	Ser mannen ham?
	'Does the man see him?'
25. POSS (No)	Knytter damen skoene hennes?
	'Is the lady tying her shoes?'
26. PP-ry (Yes)	Har mannen lagt kofferten bak seg?
	'Has the man laid the suitcase behind himself?'
27. FILLER (No)	Ser damen henne?
	'Does the lady see her?'
28. PP (No)	Har jenten satt vannkannen bak henne?
	'Has the girl put the watering can behind her?'
29. POSS-py (Yes)	Ser jenten på hatten hennes?
	'Is the girl looking at her hat?'
30. NOCTRL (No)	Holder gutten en paraply?
	'Is the boy holding an umbrella?'
31. FILLER (Yes)	Ser jenten seg selv?
	'Does the girl see herself?'
32. POSS-ry (Yes)	Løfter mannen hatten sin?
	'Is the man lifting his hat?'
33. NOCTRL (No)	Holder gutten en gris?
	'Is the boy holding a pig?'
34. PP (No)	Har mannen lagt kofferten bak ham?
	'Has the man put the suitcase behind him?'
35. FILLER (No)	Ser damen seg selv?
	'Does the lady see herself?'
36. POSS (No)	Løfter mannen hatten hans?
	'Is the man lifting his hat?'
37. PP-ry (Yes)	Har jenten satt stolen bak seg?
	'Has the girl put the chair behind herself?'
38. FILLER-py (Yes)	Ser jenten henne?
	'Does the girl see her?'
39. NOCTRL (No)	Driver mammaen og drikker en brus?
	'Is the mother drinking a soda?'
40. POSS-rn (No)	Knytter damen skoene sine?
	'Is the lady tying her shoes?'
41. PP-py (Yes)	Har gutten satt boksen bak ham?
	'Has the boy put the box behind him?'
42. FILLER (No)	Ser jenten henne?
	'Does the girl see her?'
43. PP-rn (No)	Har gutten satt bøtten bak seg?
	'Has the boy put the bucket behind himself?'

TABLE A1
Experiment 1: Number of "No" Responses per Condition per Participant

	Age	*LOCAL-py*	*LOCAL-ry*	*QLOCAL-py*	*QLOCAL-ry*	*LOCAL*	*LOCAL-rn*	*QLOCAL*	*QLOCAL-rn*
1.	5;1	0	0	0	0	3	3	3	2
2.	5;0	0	0	0	0	2	3	3	3
3.	6;0	0	0	0	0	3	3	3	3
4.	6;9	0	0	0	0	3	3	3	3
5.	5;9	1	0	0	0	3	3	3	3
6.	6;7	0	0	0	0	3	3	3	3
7.	6;7	0	0	0	0	2	3	3	3
8.	6;3	0	0	1	0	3	3	3	3
9.	5;6	0	0	0	0	2	3	2	3
10.	6;9	0	0	0	0	2	3	3	3
11.	7;2	0	0	0	0	3	3	3	3
12.	5;5	1	1	0	0	2	3	3	3
13.	6;11	0	0	0	0	3	3	3	3
14.	6;8	0	0	0	0	2	3	3	3
15.	6;10	0	0	0	0	3	3	3	3
16.	7;1	0	0	0	0	3	3	3	3
17.	6;10	0	0	0	0	3	3	3	3
18.	6;11	0	0	0	0	3	3	3	3
19.	6;5	0	0	0	0	3	3	3	3
20.	7;1	0	0	0	0	2	3	3	3
21.	6;8	0	0	0	0	2	3	3	3
22.	7;2	0	0	0	0	3	3	3	3

(Continued)

TABLE A1
(Continued)

	Age	*LOCAL-py*	*LOCAL-ry*	*QLOCAL-py*	*QLOCAL-ry*	*LOCAL*	*LOCAL-rn*	*QLOCAL*	*QLOCAL-rn*
23.	6;7	0	0	0	0	1	2	3	3
24.	6;7	0	0	0	0	3	3	3	3
25.	7;0	0	0	0	0	3	3	3	3
26.	4;5	0	0	0	0	3	3	3	3
27.	6;9	1	0	0	0	1	3	2	3
28.	5;10	0	0	0	0	3	3	3	3
29.	6;6	0	0	0	0	3	3	3	3
30.	6;3	0	0	0	0	3	3	3	3
31.	6;10	0	0	0	0	3	3	3	3
32.	5;5	0	0	0	0	3	3	3	3
33.	5;11	0	0	0	0	3	3	3	3
34.	6;8	0	0	0	0	3	3	3	3
35.	5;5	0	0	0	0	3	3	3	3
36.	5;8	0	0	0	0	3	3	3	3
37.	5;5	0	0	0	0	3	3	3	3
38.	4;6	0	0	0	1	2	3	3	3
39.	6;5	0	0	0	0	3	3	3	3
40.	6;1	0	0	0	0	3	3	3	3
41.	5;11	0	0	0	0	3	3	3	3
42.	7;4	0	0	0	0	3	3	3	3
43.	7;0	0	0	0	0	3	3	3	3
44.	5;6	0	1	1	0	3	3	3	3

TABLE A2
Experiments 2 and 4: Number of "No" Responses per Condition per Participant

	Age	*REFL-py*	*REFL-ry*	*QPP-py*	*QPP-ry*	*QPOSS-py*	*QPOSS-ry*	*REFL*	*REFL-rn*	*QPP*	*QPP-rn*	*QPOSS*	*QPOSS-rn*
2.	5;4	0	0	0	0	0	0	2	3	3	3	0	2
3.	6;4	3	0	2	0	1	0	0	3	0	3	0	3
8.	6;7	0	0	0	0	0	0	3	3	3	3	3	3
9.	5;10	0	0	0	1	0	0	3	3	1	3	2	3
12.	5;9	0	0	0	0	0	0	0	3	0	3	0	3
26.	4;9	1	0	0	0	1	0	3	2	3	2	3	3
33.	6;3	0	0	2	0	2	0	3	3	2	3	3	3
35.	5;9	0	0	0	0	0	0	1	1	2	0	2	2
36.	6;0	0	0	0	0	0	0	2	3	2	3	2	3
37.	5;9	0	0	0	0	0	0	3	3	3	3	3	3
39.	6;9	0	0	0	0	0	0	2	3	0	0	0	1
45.	6;5	0	0	0	0	0	0	3	3	3	3	3	3
46.	6;6	0	0	0	0	0	0	1	3	2	3	1	3
47.	6;7	0	0	0	0	0	0	3	3	3	3	3	3
48.	6;3	0	0	0	0	0	0	3	3	3	3	3	3
49.	6;6	—	—	—	—	—	—	—	—	—	—	—	—
50.	5;10	0	0	0	0	0	0	3	3	3	3	3	3
51.	6;7	0	0	0	0	0	0	1	3	1	3	2	3
52.	6;6	0	0	0	0	0	0	3	3	2	3	3	3
53.	5;9	—	—	—	—	—	—	—	—	—	—	—	—

(Continued)

TABLE A2
(Continued)

	Age	*REFL-py*	*REFL-ry*	*QPP-py*	*QPP-ry*	*QPOSS-py*	*QPOSS-ry*	*REFL*	*REFL-rn*	*QPP*	*QPP-rn*	*QPOSS*	*QPOSS-rn*
54.	6;0	0	0	0	0	0	0	3	3	3	3	3	3
55.	6;1	—	—	—	—	—	—	—	—	—	—	—	—
56.	6;0	0	0	2	0	0	0	0	3	0	3	1	3
57.	5;4	1	0	0	0	0	0	1	3	3	3	1	3
58.	5;8	1	0	0	0	1	0	0	3	1	3	0	3
59.	6;5	0	0	0	0	0	0	3	3	3	3	3	3
60.	5;0	0	0	0	0	1	0	1	3	1	2	3	3
61.	4;11	0	0	0	0	1	0	3	3	2	3	2	3
62.	6;0	—	—	—	—	—	—	—	—	—	—	—	—
63.	4;9	—	—	—	—	—	—	—	—	—	—	—	—
64.	6;5	—	—	—	—	—	—	—	—	—	—	—	—
65.	3;11	—	—	—	—	—	—	—	—	—	—	—	—
66.	6;5	0	0	0	0	1	0	3	3	3	3	3	3
67.	4;4	—	—	—	—	—	—	—	—	—	—	—	—
68.	6;0	0	0	0	0	0	0	3	3	2	3	3	3
69.	6;6	0	0	1	0	1	0	2	2	1	2	1	2
70.	5;0	0	0	0	0	0	0	3	3	3	3	3	3
71.	6;6	0	0	3	1	3	0	0	2	1	3	1	2
72.	6;4	0	0	0	0	0	0	3	3	2	3	3	3
73.	4;10	0	0	0	0	0	0	0	2	1	0	0	2
74.	6;8	0	0	0	0	0	0	3	3	3	3	3	3

Note. Dashes indicate missing data.

TABLE A3
Experiment 3: Number of "No" Responses per Condition per Participant

	Age	*PP-py*	*PP-ry*	*POSS-py*	*POSS-ry*	*PP*	*PP-rn*	*POSS*	*POSS-rn*
2.	5;4	0	0	1	0	0	3	2	3
3.	6;4	3	0	2	0	0	3	0	3
8.	6;7	0	0	0	0	3	3	3	3
9.	5;10	2	0	1	0	0	3	1	3
12.	5;9	3	0	2	0	0	3	0	3
26.	4;9	0	0	0	0	2	3	2	1
33.	6;3	0	0	0	0	2	3	0	3
35.	5;9	2	0	0	0	0	2	2	3
36.	6;0	1	0	1	0	1	3	2	3
37.	5;9	1	0	1	0	2	3	2	3
39.	6;9	0	0	0	0	3	0	1	1
45.	6;5	1	0	1	0	0	3	1	2
46.	6;6	3	0	1	0	0	3	1	3
47.	6;7	0	0	0	1	3	1	2	3
48.	6;3	3	1	0	0	3	3	2	3
49.	6;6	3	0	0	0	0	3	1	3
50.	5;10	0	0	1	0	2	3	2	3
51.	6;7	0	0	0	0	1	3	2	3
52.	6;6	2	0	1	0	1	3	1	3
53.	5;9	0	0	1	0	1	2	1	2
54.	6;0	1	1	1	0	2	3	2	3
55.	6;1	0	0	0	0	2	3	2	3
56.	6;0	3	0	1	0	0	3	0	3
57.	5;4	1	0	2	0	0	3	0	3
58.	5;8	3	0	2	0	1	3	1	3
59.	6;5	2	0	1	0	0	3	2	3
60.	5;0	1	0	1	1	0	3	0	3
61.	4;11	2	0	1	0	1	2	2	2
62.	6;0	0	0	0	0	0	0	0	0
63.	4;9	2	0	0	0	1	3	0	2
64.	6;5	1	0	1	0	0	3	0	2
65.	3;11	0	0	0	0	2	2	1	3
66.	6;5	0	0	0	0	2	3	3	3
67.	4;4	0	0	0	0	3	3	2	3
68.	6;0	—	—	—	—	—	—	—	—
69.	6;6	—	—	—	—	—	—	—	—
70.	5;0	—	—	—	—	—	—	—	—
71.	6;6	3	1	0	0	0	3	0	1
72.	6;4	1	0	0	0	1	3	2	3
73.	4;10	1	0	0	0	0	1	0	1
74.	6;8	0	0	0	0	2	3	3	3

Note. Dashes indicate missing data.