Subject-Object Asymmetry in Child Comprehension of *Wh*-Questions

William Philip, Peter Coopmans,
Wouter van Atteveldt and Matthijs van der Meer
UiL-OTS and University College Utrecht, Utrecht University

1. Introduction

Prior research on the L1 acquisition of English suggests that long after children have acquired knowledge of the basic syntax of *who* questions they nonetheless have greater difficulty accessing or using this knowledge in the case of object *who* questions (O-WH) such as (1b) than in the case of subject *who* questions (S-WH) such as (1a).

(1) a. *Who is helping the boy?*

b. *Who is the boy helping?*

The evidence of this is four-fold. First, corpus studies report that S-WH occur more frequently than O-WH in spontaneous speech—even though the latter may be acquired earlier (e.g. Stromswold 1995). Second, experimental production studies report that well-formed S-WH occur more frequently than well-formed 0-WH in elicited speech (Ervin-Tripp 1970, Wilhelm & Hanna 1992, Yoshinaga 1996). Third, comprehension studies report that O-WH are misunderstood more often than S-WH under experimental conditions (Ervin-Tripp 1970, Tyack and Ingram 1977; but see Cairns & Hsu 1978). Finally, both in elicited production tasks (Wilhelm & Hanna 1992) and in comprehension tasks (Ervin-Tripp 1970), two typical errors observed with O-WH are, respectively, the elicitation of an S-WH when an O-WH interpretation seems to be intended by the child and the mis-assignment by the child of an S-WH interpretation to an O-WH test sentence. To our knowledge, the converse of this error type has never been observed for S-WH.

One possible interpretation of this collection of seemingly related findings—mentioned by Stromswold (1995:17) and discussed in O'Grady (1997:135-138)—is that they reflect a processing effect of the syntactic 'distance' between the *wh* filler and its gap. According to this SYNTACTIC DISTANCE HYPOTHESIS, the observed Subject/object asymmetry in frequency of comprehension errors, as well as the typical error of mis-assigning an S-WH interpretation to an O-WH, are direct performance effects of the length of the A-bar chain in the case of O-WH. The effects observed in natural and elicited speech production are, in turn, attributable to a strategy of

avoiding harder-to-process O-WH whenever possible—i.e. to a 'preference' for easier-to-process S-WH.

The syntactic distance hypothesis gains support from two independent observations. First, in on-line studies, normal adults are also observed to have greater difficulty processing object gap relative clauses than subject gap relative clauses (e.g. Wanner & Maratsos 1978, Frazier et al. 1983). Second, preschool children have also been observed to have more difficulty correctly producing and comprehending *what*-questions such as (2b) than *what*-questions such as (2a) (Hildebrand 1987). As O'Grady (1997) notes, this latter finding can also be interpreted as a performance effect of the syntactic complexity of the filler-gap dependency.

(2) a. *What did the little girl* [$_{VP}$ *hit* __ *with a block today*]*?*

b. *What did the little boy* [$_{VP}$ *hit a stone* [$_{PP}$ *with* __] *today*]*?*

There are, however, alternative possible explanations of the subject/object asymmetry in child performance with *who* questions such as (1). First, in at least some of the experimental studies, the findings may be artifactual due to a failure to present the target input felicitously. In Tyack & Ingram (1977), for example, the task was to answer questions such as (1b) about, say, a boy who is not helping anyone. Such use of a *wh*-question necessarily causes a minor presupposition failure, since (1b) presupposes, precisely, that there is someone who the boy is helping. (A speaker not already making this assumption would normally ask *Is the boy helping anyone?*) A child's abnormal response under such conditions, then, may simply indicate a lack of adult-like expertise at accommodating infelicitous utterances.

Another possible explanation of the special difficulty children seem to have with O-WH such as (1b) is the speculation—noted in Ervin-Tripp (1977:90)— that children's acquisition of the lexical knowledge that *who* is [+animate] may interact with a comprehension parsing strategy of always initially taking *who* to be the subject of a sentence (since subjects tend to denote animate individuals). This LEXICAL HYPOTHESIS may be articulated more concretely as follows: on hearing the word *who*, the child using this strategy first builds a partial structural analysis of the input string that defines *who* as the local grammatical subject and only then attempts to assign subsequent elements of the string to positions in this structural analysis. With S-WH, this comprehension strategy would facilitate a correct parse, but with O-WH it would create a minor garden-path, necessitating re-analysis. This would require extra processing and hence might lead to errors.

The lexical hypothesis gains some support from the observation of a possible minor developmental stage reported in Ervin-Tripp (1977). In this longitudinal elicited production and comprehension study, Ervin-Tripp notes that, before the age of 3;0, "in the questions about the object, the children give appropriate responses all of the time, if they reply...[whereas] after 3;0 an upsurge of errors begins, the proportion giving subject replies to object questions being very high around 3;10." (p.

90, author's emphasis) If this observation of a sudden shift from adult-like to nonadult-like performance truly represents a developmental stage, it lends support to the lexical hypothesis under the additional, a priori plausible, assumption that it is precisely the child's acquisition of the lexical knowledge that *who* is [+animate] that triggers the radical change in performance. The comprehension parsing strategy of defining *who* as a local grammatical subject arises when the new lexical knowledge interacts with a pre-existing comprehension strategy of taking the first [+animate] NP encountered to be the local grammatical subject.

A problem for the lexical hypothesis, though, is that it is not clear how it can account for the observation that in natural and elicited speech children also tend to avoid producing O-WH. (The syntactic distance hypothesis captures this in a straightforward manner by positing that an avoidance strategy arises in response to an intrinsic processing difficulty that affects both production and comprehension.) To cover the production facts, one might consider reformulating the lexical hypothesis as a claim that children initially mis-analyze *who* as an interrogative pronoun lexically specified to contain a nominative case feature—i.e. as a kind of interrogative analog of *he*, *she*, and *they*. However, while theoretically intriguing, such a radical claim seems hard to maintain empirically in light of (i) the abundance of corrective positive evidence in the input, (ii) the fact that most of the time children show fully adult-like grammatical performance with O-WH, and (iii) the fact that many children actually produce their first grammatical O-WH before they produce their first grammatical S-WH.

Finally, note that the lexical hypothesis makes a testable prediction, at least under the concrete formulation of it described above. It predicts that children will not show any subject/object asymmetry in their comprehension of long-distance *wh*-questions such as (3). The initial assignment of a local grammatical subject function to *who* would cause the same garden path effect in both (3a) or (3b), since ultimately *who* must find its gap in the embedded clause.

(3) a. *Who did you say* [___ *helped the boy*]*?*

b. *Who did you say* [*the boy helped* ___]*?*

A third possible explanation of children's difficulty with O-WH that must be considered is the speculation that this difficulty is a processing effect following from basic English word order. On this view, which we will call the SVO HYPOTHESIS, English children have greater difficulty processing O-WH simply because these *wh*-questions have non-canonical surface word order, as illustrated in (4). O-WH such as (4b) give rise to more comprehension errors than S-WH such as (4a), and are often mis-assigned S-WH readings, because the child parser adopts the general strategy of always assigning SVO structure as the first parse of any clause. As with the lexical hypothesis, this leads to garden path effects in the case of O-WH, but not in the case of S-WH.

(4) a. Who is helping the boy?
S V O

b. Who is the boy helping?
O S V

Unlike the lexical hypothesis, the SVO hypothesis also has a simple account of why the subject/object asymmetry also occurs in natural and elicited child productions: the child parser avoids producing O-WH because, in general, its mechanisms for generating OSV sentences are as yet weakly established. Like the lexical hypothesis, the SVO hypothesis also makes the testable prediction that the subject/object asymmetry will disappear with sentences such as (3) above, because a garden path effect would occur just as often with S-WH as with O-WH.

Summing up, prior research has identified an interesting subject/object asymmetry in children's processing of *wh*-questions but has not shed much light on the mechanisms underlying this performance anomaly. Questions remain due to gaps in the research. One problem is that prior research systematically contrasting S-WH and O-WH has restricted its attention to single-clause *wh*-questions such as (1) and (2). An examination of children's performance with long-distance *wh*-questions such as (3) is needed to test the different predictions of alternative hypotheses about the underlying cause of the subject/object asymmetry.

Another problem is that prior research has restricted its attention to English child language. An examination of other child languages is needed to determine whether the Subject/object processing asymmetry is specific to child languages like English which have a canonical SVO surface word order, as the SVO hypothesis claims.

2. A Dutch Study

In an attempt to close these gaps in research on the matter, we carried out an experimental study of Dutch children's comprehension of the long-distance *wh*-question in (5a). The surface form in (5a) is structurally ambiguous, deriving either from the underlying structure represented in (5b), which yields an S-WH interpretation, or from that represented in (5c), which yields an O-WH interpretation. Note that the Dutch *wh*-word *wie* 'who' is lexically specified as obligatorily [+animate], just like English.

(5) a. *Wie zei je dat de beer natspoot*?
who said you that the bear wet-squirted

b. Wie zei je dat [___ [de beer natspoot]] (S-WH)
S O V

c. Wie zei je dat [de beer [___ natspoot]] (O-WH)
S O V

The structural ambiguity of (5a) instantiates a very general property of Dutch which distinguishes it typologically from English: Dutch freely allows both SVO and OVS surface word orders in root clauses—especially in the case of *wh*-questions—and has canonical SOV surface word order in embedded clauses. With such input, it is hard to see how the child acquiring Dutch as an L1 could ever derive an SVO strategy for processing *wh*-questions from surface forms. In addition, as noted above, with long-distance *wh*-questions, such a strategy would have the same effect on the processing of an S-WH as on the processing of an O-WH. For two independent reasons, then, the SVO hypothesis predicts that Dutch children will show no subject/object asymmetry whatsoever in their performance with the long-distance *wh*-question in (5a). The lexical hypothesis makes the same prediction, albeit for different reasons.

In contrast, the syntactic distance hypothesis makes a very different prediction, namely that Dutch children will show a significant preference for an S-WH analysis of (5a), avoiding an O-WH interpretation due to its greater processing difficulty. More specifically, the syntactic distance hypothesis makes the following prediction: if a group of normal, roughly same-age, Dutch native speaker preschoolers are each given only one opportunity to assign a meaning to (5a), a significantly greater number of them will assign the structural analysis represented in (5b) than will assign the structural analysis in (5c), whereas a control group of adult Dutch native speakers tested under the same conditions will choose (5c) as the structural analysis of (5a) roughly as often as (5b).

In addition, we tested the *wh*-question in (6a), which has the two syntactically possible analyses in (6b) and (6c), to determine whether the predicted preference for an S-WH analysis of (5a), if it occurred, was specific to children.

(6) a. *Wie zei je dat een olifant natspoot*?
who said you that an elephant wet-squirted

b. Wie zei je dat [___ [een olifant natspoot]] (S-WH)
S O V

c. Wie zei je dat [een olifant [___ natspoot]] (O-WH)
S O V

In contexts in which a specific interpretation of the indefinite NP *een olifant* would be pragmatically inappropriate, Dutch adults should have considerable difficulty assigning an O-WH structural analysis to (6a). This is because Dutch grammar does not allow an interpretation of (6c) in which *een olifant* is non-specific, due to a general prohibition against nonspecific indefinite subjects. Thus, substantial accommodation would be needed to assign this structural analysis under these

conditions. If Dutch adults show significantly different performance with (6a) and (5a), but normal Dutch native speaker children do not, then the child performance—whatever it may be—must reflect some respect in which child cognition differs from adult cognition.

2.1. Subjects

66 monolingual Dutch preschool children and 20 Dutch native speaker adults completed the study. The children consisted of 34 boys and 32 girls and ranged in age from 4;1 to 6;9 (mean age 5;6). The adults ranged in age from 17 to 22 years (mean age 20).

The child sample consisted of two sub-groups, "Group A" (n = 26) and "Group B" (n = 40), each group tested on a slightly different set of materials (see below). It was also divided into two arbitrarily defined chronological age groups, "Older Kids" (n = 41; age range 5;5 to 6;9; mean age 5;11) and "Younger Kids" (n = 25; age range 4;1 to 5;4; mean age 4;11). Independently, the children were also divided into two groups, the "Passers" and the "Failers", on the basis of whether or not they correctly gave adult-like responses for two experimental items controlling for attention (see below). There were 47 Passers (23 girls and 24 boys; age range: 4;3 to 6;9; mean age 5;7) and 19 Failers (10 boys and 9 girls; age range 4;1 to 6;7, mean age 5;4).

2.2. Procedure, Design and Materials

The experiment made use of a PICTURE STORY GUESSING GAME PARADIGM to hide the true experimental task and to enhance the felicity of the test input. The picture story guessing game is a truth-value judgement task in which a story, with accompanying pictures, is told once and then a puppet, who cannot see the pictures but who has been listening to the story, makes "guesses" about the story. Within this paradigm, the puppet's guesses are the target input, and the child's task is to determine whether or not they are correct (true or false of the story) and reward the puppet when they are. However, in this experiment the real target input were clarification questions that the puppet occasionally asked as the story was being told. The true experimental task, then, was to answer these clarification questions.

This procedure was carried out by two experimenters, both native speakers of Dutch trained in research methodology. One experimenter, the "story-teller", sat beside the child, manipulated the pictures, and told the story. The other experimenter, the "puppet master", sat opposite the child and the story-teller and controlled the puppet, reading off its lines from a hidden script. Throughout the experiment, the pictures were hidden from the puppet and the puppet master. This made the puppet's clarification questions natural and highly felicitous. The story-teller read the story from text written in large font beside or below the pictures, as if reading from a children's picture book. The pictures were page-size color drawings.

As in a standard use of the picture story guessing game, the child was instructed, and trained in a pretest, in the task of evaluating the correctness of the puppet's

guesses. In addition, the child was told that the puppet was allowed to ask questions before making a guess and that, if this happened, she should "help the puppet out" by giving him the correct answer to his question.

While the clarification questions were the true test input, the puppet's guesses also served a function in the experiment: they were control items for attentiveness. The puppet made two guesses about the story, "CT" and "CF". CT was a correct guess and CF an incorrect one. To be grouped with the Passers, a child had to correctly judge CT true of the story and CF false of it. The Failers were children who either judged CT false, or CF true, or both.

The test materials consisted of a single story containing three clarification questions—a filler item, the test input for the "DEF" condition, i.e. (5a), and the test input for the "INDEF" condition, i.e. (6a)—and the two guesses by the puppet, CT and CF. This story is illustrated in English in Figure 1 (test input in Dutch). In the actual experiment, the pictures showed an ape wearing a red sweater (the "red ape") and an ape wearing a blue sweater (the "blue ape"). Because the reproductions in Figure 1 are in black and white, we have darkened the red ape's sweater and renamed him the "grey ape", and renamed the blue ape the "white ape". In addition, in Figure 1, the target input of each experimental condition is boldfaced and identified in parentheses, and the way in which responses to the test input were coded—S-WH or O-WH analysis—is indicated in square brackets. To control for a possible "last-seen-event" bias, the story shown in Figure 1 was used for one set of materials, that presented to the Group A children, while, for the Group B children, the same story was used but with an inverted ordering of the 2nd and 3rd scenes and of the 5th and 6th scenes. The test story shown in Figure 1 was embedded in a set of 9 other stories, always occurring as the 5th story. The other stories, which functioned as filler material, were the materials for other experiments examining the comprehension of (i) the Dutch reflexive anaphor *zich,* (ii) the Dutch universal quantifier *iedere* 'every/any', and (iii) ambiguous long-distance adjunct *wh*-questions with *waar* 'where'. The children were tested individually in a quiet corner of the preschool they were attending. The adults, who were told they were control subjects for an experiment designed for children, were also tested individually, but only by one experimenter, who presented both the test and control input directly as test questions.

Figure 1: Materials for Group A

This story is about 2 apes and a bear. They are watering a plant. One ape has a grey sweater; the other has a white sweater.

Puppet: *What's the bear wearing?* (filler clarification question)
Child: *A yellow sweater.*
Puppet: *Okay. I know it.* ***Every animal is watering a plant.*** (CT)
Child: *That's right.*

2nd scene

Suddenly the bear has an idea. He turns and squirts the white ape.

3rd scene

The grey ape finds that amusing so he quirts the bear. The white ape laughs a lot.

Puppet: *Wait a minute. I'm getting confused.*
Wie zei je dat de beer natspoot? (DEF)

Child: *The grey ape.* [S-WH]
The white ape [O-WH]

4th scene

Then 3 elephants come over. They had been watching and saw what happened.

5th scene

Then an elephant gets some water with his trunk and squirts the grey ape.

6th scene

Then the white ape squirts one of the elephants.

Puppet: *Wait, wait. Not so fast! I'm all mixed up.*
Wie zei je dat een olifant natspoot? (INDEF)
Child: *The white ape.* [S-WH]
The grey ape. [O-WH]
Puppet: *Okay, I got it.* ***Every elephant got squirted.*** (CF)
Child: *No. Only 1 elephant got squirted.*

2.3. Results

For neither the 47 Passers, nor for the 19 Failers, nor for the 66 Passers and Failers combined, did the performance of the Group A children under the DEF condition differ significantly from that of the Group B children (Kruskal-Wallis tests with H adjusted for ties; $p \leq 0.4206$, $p \leq 0.2563$, and $p \leq 0.2485$, respectively), and the same was true of the performance of these groups under the INDEF condition ($p \leq 0.7205$, $p \leq 0.6086$, and $p \leq 0.4563$, respectively). Evidently there was no significant "last-seen-event" bias, so the Group A and B subjects have been pooled in the subsequent analyses.

For the 47 Passers and 66 Passers and Failers combined, there were no significant contrasts in performance between the Younger Kids and the Older Kids, neither under the DEF condition ($p \leq 0.7840$ and $p \leq 0.1841$, respectively), nor under the INDEF condition ($p \leq 0.9880$ and $p \leq 0.8002$, respectively) . For the 19 Failers alone, there was no significant age effect for the INDEF condition ($p \leq 0.6086$), but there was for

the DEF condition ($p \leq 0.0029$). This latter finding, however, seems to be a statistical artifact and/or an effect of the inattentiveness of the Failers who were Older Kids: while only 3 of the 8 Failers who were Younger Kids chose the S-WH interpretation under the DEF condition, all 11 of the Failers who were Older Kids gave this response. Taken together, these statistical analyses suggest that age was not a relevant factor in this experiment, so we will collapse the two chronological age groups in subsequent analyses.

The Failers consisted of 14 children who judged CF true and 5 who judged CT false. Their performance under the DEF and INDEF conditions did not contrast significantly with that of the Passers, but below we will keep these groups separate nonetheless, for clarity. The average percentages of S-WH interpretations of the two test input sentences are shown in Table 1 (standard errors indicated in parentheses). The rightmost column shows the significance of the respective DEF/INDEF contrast as determined by a sign test.

Table 1: Average Percentages of S-WH Interpretations

groups	n	DEF	INDEF	sign tests
Adults	20	45% (11%)	80% (9%)	$p \leq 0.0391$
All Children	66	70% (6%)	74% (5%)	$p \leq 0.6892$
Passers	47	68% (7%)	77% (6%)	$p \leq 0.4545$
Failers	19	74% (10%)	68% (11%)	$p \leq 1.0000$

The performance of the 66 children as a whole contrasted significantly with that of the 20 adults under the DEF condition (Kruskal-Wallis with H adjusted for ties; $p \leq 0.0451$), but not under the INDEF condition ($p \leq 0.6017$). Comparing the Passers with the adults, the significance of the contrast in performance under the DEF condition was slightly less significant ($p \leq 0.0782$).

4. Conclusion

Our results support the syntactic distance hypothesis and argue against the lexical and SVO hypotheses. The prediction of the former hypothesis was borne out; that of the latter two was not. It seems that it is a universal property of language processing that the further a *wh*-expression is removed from its gap the more difficult it is to process *wh*-movement. The effects of this difficulty are so pronounced in preschool children's performance that they even can be detected with off-line techniques.

References

Cairns, Helen and Jennifer Ryan Hsu (1978) "*Who, Why, When*, and *How*: A Developmental Study", *Journal of Child Language* 5, 477-488.

Ervin-Tripp, Susan (1970) "Discourse Agreement: How Children Answer Questions", in *Cognition and the Development of Language*, ed. John R. Hayes, 79-106. Wiley: New York.

Frazier, L., C.Clifton, and J. Randall (1983) "Filling Gaps: Decision Principles and Structure in Sentence Comprehension", *Cognition* 13, 187-222.

Hildebrand, Joyce (1987) "The Acquisition of Preposition Stranding", *Canadian Journal of Linguistics* 32, 65-85.

O'Grady, William (1997) *Syntactic Development*. University of Chicago Press: Chicago, Illinois.

Stromswold, Karin (1995) "The Acquisition of Subject and Object *Wh*-Questions", *Language Acquisition* 4, 5-48.

Tyack, Dorothy and David Ingram (1977) "Children's Production and Comprehension of Questions", *Journal of Child Language* 4, 211-224.

Wanner, Eric and Michael Maratsos (1978) "An ATN Approach to Comprehension", in *Linguistic Theory and Psychological Reality*, eds. Moris Halle, Joan Bresnan and George A. Miller, 117-161. MIT Press: Cambridge, Massachusetts.

Wilhelm, Andrea and Ken Hanna (1992) "On the Acquisition of *WH*-questions", *Calgary Working Papers in Linquistics* 15, 89-98.

Yoshinaga, Naoko (1996) *WH Questions: A Comparative Study of Their Form and Acquisition in English and Japanese*. PhD dissertation, University of Hawaii at Manoa.