Second Language Acquisition of Reflexive Binding

Daniel L. Finer
SUNY at Stony Brook

Ellen I. Broselow
SUNY at Stony Brook

Second Language Acquisition of Reflexive Binding*

Daniel L. Finer and Ellen I. Broselow

SUNY at Stony Brook

1 Introduction

This paper reports on a pilot study designed to investigate the acquisition of English reflexives, specifically the domain of English reflexive binding, by native speakers of Korean. We were interested in three major questions: first, the question of whether and to what extent the setting of a parameter in the second language is influenced by the parameter setting of the native language; second, whether the acquisition of the second language is governed by the same principles which have been argued to govern first language acquisition; and third, the extent to which universal principles of markedness influence second language acquisition. Since English and Korean differ in the domains in which reflexives may be bound, with English at one end of the markedness hierarchy proposed by Wexler and Manzini (1985, to appear)[1] and Korean at the other end, the emerging grammar of Korean learners of English provides an interesting area of research into these questions. We begin by discussing the facts of reflexive binding in Korean and English and locating these languages on the hierarchy proposed by Wexler and Manzini. We then discuss the predictions of the hierarchy, along with the Subset Principle proposed by Berwick (1985) for first language acquisition of reflexives. Next we discuss the intuitions of the Korean learners about sentences containing English reflexives, and offer several possible interpretations of the data. We find that the emerging grammar of the Korean learners is different from both the grammar of the first language and the grammar of the second language, but it is still consistent with proposed parameters of universal grammar.

154

2 The Subset Principle and its Role in First Language Acquisition

It is well known that languages may differ with respect to the domains in which a lexical anaphor can find its antecedent (see in particular Harbert 1982, Yang 1983, etc). In English, for example, the reflexive and its antecedent obey what may loosely be called a clausemate condition, while in Korean, the antecedent for the reflexive caki may be local, or it may be several clauses away from the reflexive. Thus, Korean allows the range of interpretations in (1a), while English is restricted to the interpretation in (1b). (Here we abstract away from differences in word order.)

(1)   NP V [NP V [NP V Refl.]]

   a. i   j   k   i/j/k   (Korean)

   b. i   j   k   *i/*j/k (English)

The two languages converge, of course, in a structure like that in (2):

(2) NP_i V $Refl_i$

Restricting attention to Korean and English for now, note that English is a subset of Korean when it comes to interpretations of sentences containing reflexive anaphors: that is, for every well-formed English interpretation, there is a corresponding well-formed interpretation in Korean, but not the reverse. As it turns out, English and Korean exemplify the two extremes in terms of possibilities for defining the domain of reflexive binding; English is the most restrictive and Korean (along with Japanese) the least restrictive. In between, as noted by Manzini and Wexler, are other languages less restrictive than English but more restrictive than Korean. Wexler and Manzini note that the domain in which a reflexive in a given language must be bound is sensitive to properties of sentential inflection, and that these properties induce a set-theoretic relation of proper inclusion between various languages on well-formed interpretations of the sort illustrated above between English and Korean. They then propose that the definition of 'governing category' be associated with a five-valued parameter, as in (3):[2]

(3)   Governing Category Parameter

g is a governing category for a iff g is the minimal
category which contains a, a governor for a and has

a.   a subject; or

b.   an Infl; or

c.   a Tns; or

d.   an 'indicative' Tns; or

e.   a 'root' Tns

Languages that contain anaphors found along this hierarchy
are English for type (a), Italian for type (b), Russian for type
(c), Icelandic for type (d), and Korean and Japanese for type
(e).[3]

As an example of the subset relations sketched above,
consider the following sentence:

(4)   Keith$_i$ said that Ronnie$_j$ requires that Bill$_k$

persuade Charlie$_l$ to consider Mick$_m$ fond of himself$_X$

In a language of type (a), X can only equal m. X may
equal either l or m in a type (b) language, k, l, or m in a type
(c) language, and j, k, l, or m in a type (d) language.
Finally, in (e), the most liberal language type, X may equal any
of i, j, k, l, or m. Thus, as we move from (a) to (e) in the
above hierarchy, the range of interpretations increases in an
ever-widening relation of proper inclusion; every interpretation
higher on the hierarchy is included in the set of interpre-
tations lower in the hierarchy, but not the reverse.

Let us assume this hierarchy as a matter of linguistic
fact, and then go on to consider its relationship to questions
of learnability. Wexler and Manzini note that under the
assumption that only positive evidence is relevant to first lan-
guage acquisition, the (a) language type must be the least
marked; that is, (a) must be the initial setting of the
five-valued parameter in (3). So if a child in, for example, an
English-speaking environment hears (4) above, the (a) setting on
the domain parameter will force the child to interpret the re-
flexive with Mick as the antecedent. Given the grammar of
English, the child will never be exposed to positive evidence
that will cause the Governing Category parameter to be reset.
On the other hand, if a child is in, for example, a

Russian-speaking environment, and is exposed to (4), the child's initial assumption will again be that Mick is the antecedent. But in this case, there will eventually be positive evidence to the effect that Charlie and Bill are also possible antecedents, leading the child to reset the parameter.

Alternatively, consider the scenario if (e) were the initial setting. Here, the child's initial assumption would be that any of the names in (4) could be antecedents for the reflexive. This assumption would present no problems if the child were in an environment in which a type (e) language was spoken. However, a child in an environment in which a language of type (a), (b), (c), or (d) was spoken would require negative evidence to land on the right parameter setting, since every interpretation of (4) in the target language would be compatible with the hypothesis that (e) defines the binding domain of the reflexive (following the Subset Hierarchy).

The above considerations, along with the assumption that children do not in fact have significant access to negative evidence in the acquisition process (see for example Baker 1979), lead Manzini and Wexler to the hypothesis that the Subset Principle on learnability --- which can be informally paraphrased as saying that the learner will select the value of a parameter which generates the smallest language that is compatible with the data --- will play a significant role in the acquisition of anaphora.[4]

3 The Problem of L2 Acquisition

With this as background, then, we can now turn to the problem of second language acquisition and consider what might take place in the acquisition of reflexive binding when the native language and the target language occupy different positions on the Wexler-Manzini markedness hierarchy. The particular question we will consider is what might be expected to occur when Koreans, whose native language has the most marked setting of the parameter for reflexive binding, learn English, which has the least marked setting.

There are at least four possibilities, and two appear prima facie most plausible. First, one might expect that Korean speakers would simply transfer the parameter setting exemplified by Korean into English and would therefore make mistakes such as taking all of the Stones in (4) as possible antecedents of the reflexive. In this case, negative evidence would be crucial to

convince the Korean learners to narrow the range of possible
antecedents. Alternatively, however, one might assume that the
acquisition of the binding principles of a second language
proceeds as does the acquisition of the binding principles of
the first language --- learners start with the unmarked value
(giving Mick as the only possible antecedent in (4)), and stay
at that unmarked value until positive evidence forces them to do
otherwise. Here, one would expect few mistakes from Koreans
learning English, since these learners would presumably ignore
evidence from Korean and start fresh, as it were, with the most
restrictive hypothesis about binding domains. A third
possibility, one that seems perhaps less likely, is that the
grammar of the learners of English would exemplify the binding
principles of neither English nor Korean. This third
possibility in turn divides into two subpossibilities. First,
the grammar of the Korean learners might display the binding
properties of no natural language --- for example, it might
allow only local binding in infinitive clauses but long-distance
binding out of tensed clauses. Or second, the Koreans' grammar
might conform to the principles of neither English nor Korean
but instead to those of some other natural language, one
intermediate between the two on the hierarchy.[5] The latter
possibility is the one suggested by our results.

4 The Experiment

The pilot experiment we carried out was designed to
discover what principles did in fact govern reflexive binding in
the English of these Korean learners. The experiment involved
six subjects, all adult students in the intensive English
program at the State University of New York at Stony Brook and
newly arrived in the United States. Five of the subjects were
intermediate or advanced; one had been placed in the highest
level of a six-level program, two in level 5, and two in level
4. The sixth subject had had only two years of instruction in
English and was placed in level 2. The experiment involved a
picture identification task. Subjects were shown sets of
pictures like the ones included in Appendix A. Each set
contained two pictures, one labelled A and the other labelled B.
Subjects were asked to indicate whether the sentence which went
along with the picture was appropriate for picture A, picture B,
or both A and B. All pictures involved the characters Mr. Fat
and Mr. Thin, and subjects were first introduced to these
characters and given two sentences involving no reflexives or
pronouns at all, to make sure they understood the task. They
were then given four simple sentences of which two contained re-

flexive objects and two contained pronominal objects, followed
by sixteen complex sentences, of which half contained reflexives
and half contained pronominals. In each set, half the embedded
clauses were tensed while the other half were infinitival
clauses. Examples of the sorts of sentences used are shown in
Appendix A. These sentences would of course be ambiguous in
Korean, with the preferred interpretation being the nonlocal one
--- in contrast to English, where the only interpretation is the
one in which the reflexive has a local antecedent.

4.1 Results

        The responses of the subjects to the sixteen complex
sentences are shown in Appendix B. If we consider the reflexive
results as a whole, it is clear that the Korean subjects did not
simply apply the binding principles from their native language
to the English sentences. Instead, they showed a clear
preference for the local interpretation of the English reflex-
ives, giving thirty-six responses indicating a local interpre-
tation, eleven indicating a nonlocal interpretation, and only
one response indicating that they found the sentence ambiguous.
It also seems clear, however, that the subjects had not yet
internalized the principles which govern the distribution of re-
flexives in English, nor are they, at this stage at least,
retreating to the most restrictive grammar, since they have
still given twelve 'wrong' responses out of a total of
forty-eight.

        But an interesting pattern emerges when we look more
closely. Note first of all the rather striking differences
between those sentences involving a tensed complement compared
with those involving an infinitival complement. Over 90% of the
responses to sentences involving tensed complements followed
English binding principles. In contrast, of the responses to
sentences involving infinitival complements, less than 60% were
consistent with English binding principles. In general, it
looks as though the subjects show a tendency to follow English
principles for tensed complements and Korean principles for
infinitival complements. (The subject-by-subject breakdown
shows that while only one subject makes a real flip from local
to nonlocal interpretation on the infinitive reflexive clauses,
all subjects show a greater tendency to interpret reflexives in
tensed clauses as local more often than reflexives in infinitive
clauses.) What makes these results surprising is that the
tensed/infinitival distinction is irrelevant to the distribution
and interpretation of reflexive anaphors in both the native lan-
guage and the target language. Korean does not appear to have
either an agreement marker (AGR) or infinitival verbs; thus, one
cannot argue that a Korean distinction has been realized in two

different ways in the second language since there is no Korean distinction here. Secondly, while English does have the tensed/infinitival distinction, this distinction does not play a role in the distribution of reflexives, so it cannot be argued that the distinction shown in the data is due to the grammar of English. In short, the distinction that emerges cannot be attributed to either the native language or the target language grammar.

5 Interpretation

One possible interpretation of these results is driven by the hierarchy of binding domains discussed earlier. In terms of this hierarchy, it might be provisionally proposed that the subjects are assuming a value for the Governing Category parameter that is somewhere between the (a) value (English) and the (e) value (Korean), possibly (c) (Russian). The subjects seem to have converged on language type (c) or (d). (Since there were no subjunctive clauses in the stimuli, we cannot determine the cut-off point here.) What appears to be the case is that the subjects have essentially 'split the difference' between Korean and English, and in so doing, they have come up with a set of binding principles that is consistent with the parameters provided by universal grammar, but inconsistent with either the first or the second language.

Regardless of how the subjects arrived at this point, it is important to note that the distinction that was made by these subjects between tensed and infinitival complements is indeed consistent with the range of possible grammars allowed by the theory assumed here. A logically possible outcome would have been one in which local antecedents were chosen in infinitives and non-local antecedents were chosen in tensed clauses, as in (5):

(5) a. Mr. Fat$_i$ thinks that Mr. Thin$_j$ will paint himself$_i$.

b. Mr. Fat$_i$ wants Mr. Thin$_j$ to paint himself$_{j/*i}$.

The pattern in (5), however, is incompatible with the markedness hierarchy above, and further, no language with this binding pattern has been attested. The Wexler-Manzini hierarchy rules it out in principle, since Infl is a larger class than Tns; anything bound in a category containing Tns is also bound in a category containing Infl, but not the reverse. [6]

Thus, principles of universal grammar are clearly playing a role here. What is less clear, however, is the role of language transfer, on one hand, and of the Subset Principle, on the other. If language transfer is the dominant factor, we would expect our subjects to have started out with the Korean parameter setting and, with exposure to negative evidence --- presumably available to them in the language classroom --- to have gradually narrowed the domain of reflexive binding, moving up on the hierarchy in (3). On this interpretation, we would expect to find a correlation between the ability level of a learner and his parameter settings, with a tendency for the parameter setting to move from (e) upwards. Unfortunately, we do not have the data to test for the presence of such a tendency.

In terms of the Subset Principle, however, it appears that these results do not support the hypothesis that subjects reactivated the Subset Principle. If they did so, they should start out with the English setting, and since there would be no evidence from English to cause them to then move to a more marked setting on the parameter, we would expect them to prefer the local interpretation of reflexives in all the English sentences. On this first interpretation, then, it looks as though the Subset Principle is inapplicable to second language acquisition, or at least to the acquisition of principles of reflexive binding.

5.1 Another Interpretation and Another Parameter

There is a second analysis of the data, however, in which the Subset Principle is fully operative. Before we go into the discussion, however, it is necessary to note an additional difference between reflexives in English and Korean. Besides the difference in binding domains, the two languages differ also in their choice of what can count as a proper antecedent for the reflexive. In Korean, only subjects can be antecedents, while in English, both subjects and objects are potential antecedents. In a sentence like (6), then, English and Korean differ in terms of possible binding relationships.[7]

(6)   Wally$_j$ showed Ward$_i$ a picture of himself$_X$.

In (6), X may be equal to either i or j in English, but X may only equal j in the Korean counterpart. With respect to this difference, governed by what Wexler and Manzini call the Proper Antecedent Parameter, Korean is unmarked and English is marked --- the well-formed interpretations in Korean are a subset of the well-formed interpretations in English, and the subset relation between the two languages is reversed from that

given by the Governing Category Parameter.

So, according to the Governing Category Parameter, English is unmarked and Korean is marked, while according to the Proper Antecedent Parameter, it is Korean which is unmarked and English which is marked. The experimental results can be interpreted as conforming to the Subset Principle with respect to the Governing Category Parameter if one additional assumption is made: that subjects analyzed the subject of the infinitive as the direct object of the matrix verb --- essentially a raising to object analysis, with no trace left by movement[8] --- and they have either omitted or ignored PRO in the control structures. We should note here that no difference emerged in the results between control verbs and exceptional case-marking verbs in terms of their licensing of nonlocal antecedents.

This "raising to object" analysis is represented in (7).

(7) a.   Mr. Fat$_j$ thinks [$_S$that Mr. Thin$_i$ will paint himself$_X$]
          SUBJ   SUBJ

    b.   Mr. Fat$_j$ wants Mr. Thin$_i$ [$_{VP}$to paint himself$_X$]
          SUBJ   OBJ

Assuming that this is the analysis in terms of grammatical relations assigned by the subjects, the unmarked value of the Governing Category Parameter, taken independently, will offer only i as a value for X in (7a) and either i and j as values for X in (7b). The unmarked value of the Proper Antecedent Parameter, however, yields only i as a value of X in (7a) and only j as a value for X in (7b), since the reflexive has access only to subjects, and Mr. Thin is analyzed as an object. With the unmarked settings for both parameters, then, we would expect subjects to prefer the local interpretation on tensed clauses and the nonlocal interpretation in infinitivals. But while reflexives in infinitival clauses were interpreted as nonlocal significantly more often than reflexives in tensed clauses, the reflexives in the English infinitivals were still interpreted as coreferential with the local subject (read object) in the majority of the responses. This suggests that subjects cannot have been limiting their choice of antecedents only to subjects.

The recognition of the difference in the Proper Antecedent Parameter in English and Korean, however, suggests an approach wherein the Subset Principle plays a large role. We might assume, as above, that the Korean learners are taking the subjects of the infinitives to be objects, but that they are applying the Proper Antecedent Parameter with its English setting. According to this hypothesis, the choice of antecedent

with the infinitives is free (that is, either NP is a potential
antecedent), but the antecedent is restricted to the embedded
subject with tensed complements because there are no other NP's
within the restricted governing category. This analysis assumes
the unmarked setting for the governing category and the marked
setting for the proper antecedent, and therefore predicts that
either the matrix or the embedded subject can be the antecedent
of the reflexive in the infinitives. This analysis is perfectly
consistent with the varying interpretations of the reflexives in
the infinitival clauses.

        We might further hypothesize that although the marked
setting for the Proper Antecedent Parameter is operative, this
is not necessarily in contradiction to the Subset Principle,
since there is in principle data available in the form of
positive evidence to lead learners toward the marked setting.
That is, if the Korean learners reset both parameters to their
unmarked settings, there would be no evidence that would force
them to change the setting for the Governing Category Parameter,
but there would be evidence, in the form of sentences like (6),
that would force a resetting of the Proper Antecedent Para-
meter.

        Further experiments which would be useful in deciding
between the analyses offered above would elicit interpretations
of sentences like the one in (6), in order to determine what the
subjects would take as the antecedent of the reflexive. If they
were unwilling to take Ward (the direct object) as an antece-
dent, then the raising-to-object interpretation would lose much
of its appeal, since in the majority of the infinitival cases
the embedded subject (the putative derived object) is the ante-
cedent of the reflexive.[9]

6 Conclusion

        The results discussed above, along with the suggested
interpretations of these results, bear on several interesting
issues: the relevance of the Subset Principle to second language
acquisition, the interdependence of various parameters, and the
role of positive and negative evidence in second language
learning. The evidence is inconclusive on most of these points,
but one result emerges quite clearly: that the developing
grammars of these language learners were consistent with the
subset of logically possible systems found in actual human lan-
guages. Thus it is clear that these subjects were constrained
by the range of possibilities offered by universal grammar in

their construction of the grammar of a second language.

APPENDIX A

1. Mr. Fat (Mr. Thin) thinks (wishes/knows/believes) that Mr.
      Thin (Mr. Fat) will paint himself (him).

2. Mr. Fat (Mr. Thin) asks (tells/expects/wants) Mr. Thin (Mr.
      Fat) to paint himself (him).

Finer & Broselow

165
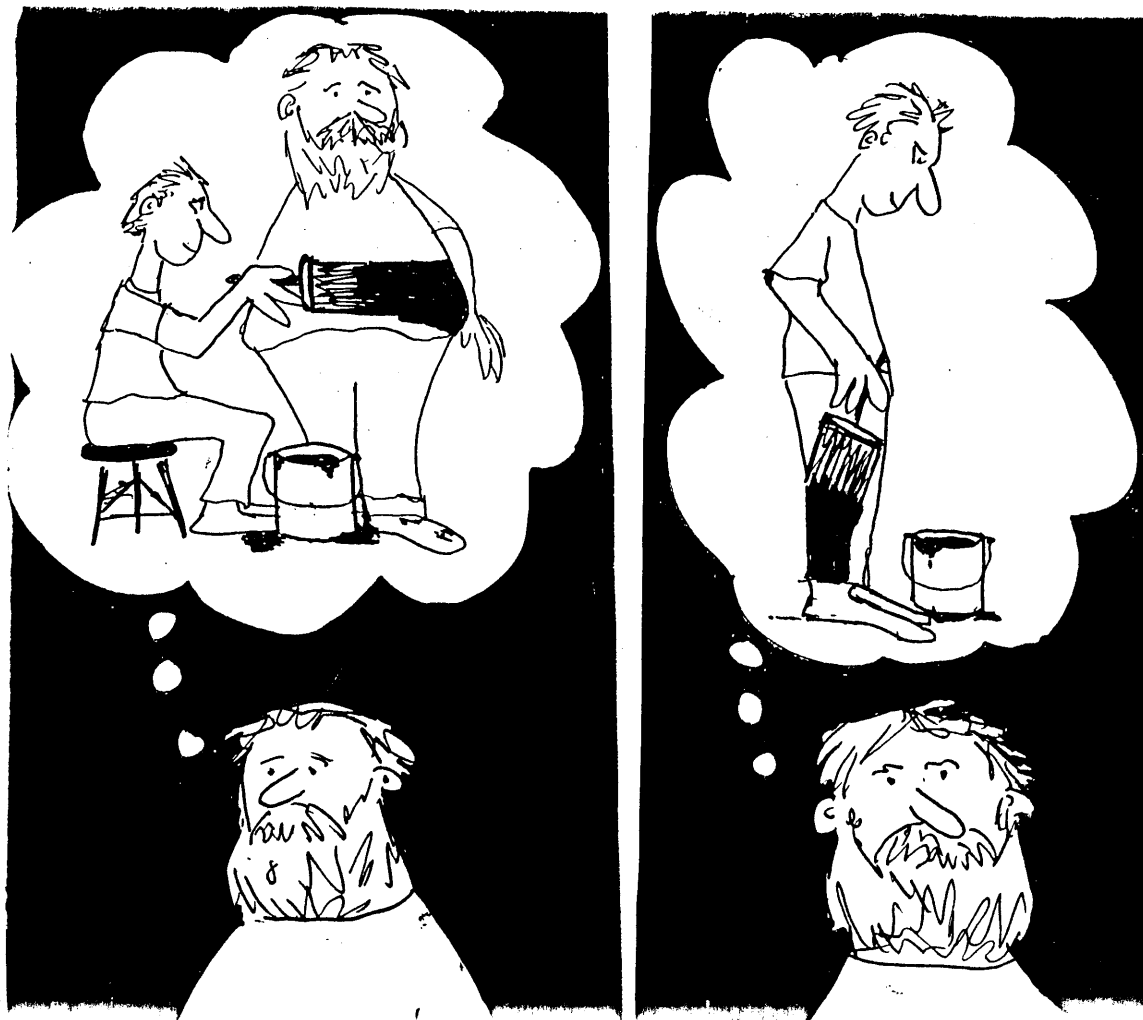
APPENDIX B

REFLEXIVES: ALL SUBJECTS

	Local	Nonlocal	Either
Tensed Clause	22	2	0
%	91.7%	8.3%	0%
Infinitive Clause	14	9	1
%	58.3%	37.5%	4.2%
Total	36	11	1

PRONOUNS: ALL SUBJECTS

	Local	Nonlocal	Either
Tensed Clause	7	13	4
%	29.2%	54.2%	16.7%
Infinitive Clause	1	19	4
%	4.2%	79.2%	16.7%
Total	8	32	8

RESPONSES BY SUBJECT

Subject (level)		Local	Nonlocal	Either
JHK (4)	TR	4	0	0
	IR	1	3	0
	TP	1	3	0
	IP	0	4	0
CES (5)	TR	4	0	0
	IR	3	1	0
	TP	2	2	0
	IP	0	4	0
HSK (6)	TR	4	0	0
	IR	3	1	0
	TP	1	3	0
	IP	1	3	0

SH (5)	TR	4	0	0
	IR	3	1	0
	TP	3	1	0
	IP	0	4	0
HKS (2)	TR	4	0	0
	IR	3	1	0
	TP	0	4	0
	IP	0	4	0
JYS (4)	TR	2	2	0
	IR	1	2	1
	TP	0	0	4
	IP	0	0	4

FOOTNOTES

*We are greatly indebted to Mee-Hwa Lee for help in designing and administering the experiment, serving as informant, and offering many insightful comments on the results. The second interpretation of the data developed from her suggestions. We are also grateful to Ken Wexler, Frank Anshen, and Mark Aronoff for comments and technical advice.

1. This and all subsequent references to Wexler and Manzini's work are to Wexler and Manzini (to appear) and Manzini and Wexler (1985)

2. A clause on accessibility is omitted here.

3. See the papers by Wexler and Manzini for further discussion.

4. See Manzini and Wexler (1985) for a formal statement of this principle.

5. None of these possibilities is to be expected any more than the others in the absence of a theory of grammar and a theory of how that grammar comes to be acquired by adults.

6. Or, in EST terms, the actual results are compatible with a view that the Specified Subject condition is suspended while the Propositional Island Condition is obeyed; it is impossible, however, to interpret the pattern in (5) in EST terms.

7. During the discussion, it was pointed out by one of the

participants that there are Korean cases in which a direct
object can serve as the antecedent of _caki_. This observation
was countered by another participant who suggested that what
looks like a direct object may be analyzed as the subject of a
small clause. There also appears to be some dialectal variation
here as well, since our informant refuses to accept any
non-subject as the antecedent of a reflexive.

8. This is a reasonable enough analysis to those familiar with
GPSG, LFG, and earlier versions of transformational grammar.

9. Another potential problem with the raising interpretation
arises when the data from the pronoun test is examined. With
the raising analysis, there would be no distinction between
local and non-local with respect to a subject antecedent --- the
matrix subject would be the only subject in the sentence --- and
in the data there was a strong tendency to take the matrix
subject as the antecedent for the pronoun, even moreso than for
the tensed complements (we have no explanation for this
contrast). So if the governing category of the embedded direct
object is the entire clause in these examples (given that there
is only one subject), we have even more Principle B violations
to account for than before. See Jakubowicz 1984, Solan (to
appear) and Wexler and Chien (in press) for discussion of
similar facts in first language acquisition, and Montalbetti and
Wexler 1985 for revision of Principle B which reflects the ac-
quisition facts.

REFERENCES

Baker, C.L. (1979) "Syntactic Theory and the Projection
    Problem," _Linguistic Inquiry_ 10:4.

Berwick, R. (1985) _The Acquisition of Syntactic Knowledge._ MIT
    Press.

Chomsky, N. (1981) _Lectures on Government and Binding._ Foris.

Harbert, W. (1982) "Should Binding Refer to SUBJECT?" NELS 12.

Jakubowicz, C. (1984) "On Markedness and Binding Principles,"
    NELS 14.

Manzini, R. and K. Wexler. (1985) "Parameters, Learnability and
    Binding Theory," ms. UC Irvine.

Montalbetti, M. and K. Wexler (1985) "Binding is Linking," WCCFL
    IV.

Solan, L. (to appear) "The Acquisition of Pronouns and
    Reflexives," in T. Roeper and E. Williams, eds. _Parameters
    and Linguistic Theory_, Reidel.

Wexler, K. and Y.C. Chien. (in press) "The Development of
    Lexical Anaphors and Pronouns," in _Proceedings of the 1985
    Child Language Research Forum_.

Wexler, K. and R. Manzini. (to appear) "Parameters and
    Learnability in Binding Theory," in T. Roeper and E. Williams,
    eds. _Parameters and Linguistic Theory_, Reidel.

Yang, D. W. (1983) "The Extended Binding Theory of Anaphors,"
    Language Research 19:2.