

1984

On Markedness and Binding Principles

Celia Jakubowicz
MIT, CNRS

Follow this and additional works at: <https://scholarworks.umass.edu/nels>



Part of the [Linguistics Commons](#)

Recommended Citation

Jakubowicz, Celia (1984) "On Markedness and Binding Principles," *North East Linguistics Society*. Vol. 14 , Article 11.

Available at: <https://scholarworks.umass.edu/nels/vol14/iss1/11>

This Article is brought to you for free and open access by the Graduate Linguistics Students Association (GLSA) at ScholarWorks@UMass Amherst. It has been accepted for inclusion in North East Linguistics Society by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

ON MARKEDNESS AND BINDING PRINCIPLES*

CELIA JAKUBOWICZ

MIT Center for Cognitive Science
 Laboratoire de Psychologie Expérimentale, CNRS, Paris

Linguistic research, addressing the problem of language acquisition, has conceived of acquisition as an instantaneous process (Chomsky (1965)) in which "principles" and "data" are massed together in an attempt to arrive at an intelligible idealization. In fact, acquisition is not instantaneous. Language, like many other biological "organs" obeys a regular maturational course. By adopting an explicit developmental approach which allows us to make the instantaneous model consistent with the non-instantaneous child, we attempt to answer certain questions concerning the setting of parameters and language growth. More precisely, the present study examines certain issues concerning binding properties of overt anaphors and pronouns within the general framework of the Government and Binding Theory (Chomsky (1981)) and uses as evidence, data from experimental research on the comprehension of these expressions by young English-speaking children. We conclude that a young child's interpretation of referentially dependent NPs is constrained by a principle of local A-binding.

1. Binding and Locality

A considerable amount of research has been devoted to the study of overt anaphors and pronouns in order to find the invariants of the anaphoric binding system of

CELIA JAKUBOWICZ

natural languages. In the Government and Binding Framework outlined in Chomsky (1981, 1982), nominal expressions are divided into three basic categories: (i) anaphors (reciprocals, reflexives, NP traces, PRO), (ii) pronouns (PRO, phonetically realized pronouns), and (iii) R-expressions (names, variables). The distribution of these nominal expressions is constrained by the Binding Principles.

(1) Binding Principles

- A. An anaphor is A-bound in its governing category.
- B. A pronominal is A-free in its governing category.
- C. An R-expression is free.

Informally speaking, these principles require anaphors to be c-commanded by an antecedent in an argument position (A-position) in a given domain. In the same domain, pronouns must be free; i.e. must not be c-commanded by an antecedent in an A-position. Finally, R-expressions must be free; i.e. not have any antecedent. The domain in which an anaphoric expression must be bound and a pronominal free is defined by the Binding Theory in terms of governing category, namely,

- (2) β is the governing category for α if and only if β is the minimal category containing α , a governor for α , and a SUBJECT accessible to α .

"Accessibility" is defined in (4), in terms of the well-formedness condition given in (3).

- (3) $*[\gamma \dots \delta \dots]$, where γ and δ bear the same index.
- (4) α is accessible to β if and only if β is in the c-commanding domain of α and assignment to β of the index of α would not violate (3).²

Chomsky (1981) remarks that the formulation above focuses attention on a phenomenon that has only been partially captured by previous approaches: the near complementary distribution of pronouns and reflexives. However, it has been shown that in Dutch, Icelandic, Japanese and many other languages the element translated as Reflexive in English does not observe the Binding Theory as sketched above. A distinction has been made between not locally bound anaphors, which may appear in free variation with pronouns, and locally bound anaphors which may not. Hence, the reflexive/non-reflexive complementarity has

ON MARKEDNESS AND BINDING PRINCIPLES

been denied the status of an invariant of the anaphoric binding system. Notice that the above distinction concerns reflexive anaphors but not reciprocals. The latter are the anaphors with the least variation across languages and their behaviour can be captured by the Principles of Binding.

With respect to reflexive anaphors, cross-linguistic data show that the situation is more complex. On the one hand, a distinction has to be made between languages such as English, which possess only one form of reflexive anaphor, and languages such as Italian or Norwegian, where several morphological realizations appear. On the other hand, languages may vary according to whether their grammars license both local and not local binding for some reflexive element(s) (cf. Icelandic) or whether the grammar requires particular reflexive elements to be specified either for local or for not local binding (cf. Norwegian). Thus, the distribution of reflexives across languages appears to be a complex matter.

However, a more careful analysis of some basic cross-linguistic facts does allow us to argue that some of the apparently drastic differences in anaphor binding across languages follow naturally from Chomsky's framework, given an adequate parameterization of the domain of binding. Let us then assume that the domain of binding is the parameter associated with the Binding Principles and that, in the unmarked case, this domain has to meet a strict locality requirement. In other words, in the unmarked case, an accessible SUBJECT (subject or AGR), creates an opaque domain. Further, Universal Grammar (henceforth UG) leaves the parameter with an open option, allowed by conditions which vary across languages. The open option consists of skipping a specified subject. From a descriptive point of view, let us then consider that reflexives falling under this option form instances of non-local binding.⁴

To illustrate the above, consider Icelandic reflexive sig, which, as mentioned by many authors, may undergo long distance binding when it appears in a subjunctive clause. Otherwise, it is locally bound. (Sig is a third person reflexive, invariant for gender and number; it has accusative, dative and genitive forms.)⁵ Examine the following sentences.

- (5) a. Jon_i segir að María elski sig_i
 Jon says that María loves/SUBJ self
- b. *Jon_i segir að María elskar sig_i
 Jon says that María loves/IND self

CELIA JAKUBOWICZ

The examples above illustrate the fact that it is only possible for sig not to be locally bound when it appears in a clause where the tense element is subjunctive (5a). When the tense element is indicative (5b), the embedded clause is the governing category of sig, which then must be bound; coindexing sig with Jon in the matrix clause violates Principle A, hence (5b) is ungrammatical.

The same generalization holds in Italian, where se (=him/herself) and proprio (=his/anaphoric) may be not locally bound when they appear in a subjunctive embedded clause. (Giorgi, 1983)⁶

Turning now to Norwegian, we find still another pattern of reflexive distribution. Norwegian possesses two kinds of reflexive NPs: reflexive NPs containing seg (=him/herself) and reflexive NPs containing selv (seg selv). Selv, in its reflexive use, occurs either with seg or the non-reflexive pronouns ham, henne, (=him, her). Seg and seg selv may have masculine or feminine, singular or plural antecedents, but not first or second person antecedents. The data below, described in Hellan (1980) show that in Norwegian, the reflexive elements are in complementary distribution. Roughly speaking, seg selv and ham selv must be locally bound, the former to a subject, the latter to an object. Seg is specified for long distance binding only, but crucially this property is licensed when the tense element of the clause in which seg appears is an infinitival. Hence, in sentence (10) below, where a free tense occurs in the embedded clause, long distance binding is blocked: coindexing of seg and Ola is ungrammatical. For seg selv, in the embedded clause, the AGR element of the INFL node is accessible to seg selv. Thus, the embedded clause is the governing category for seg selv, which must be bound. But in this domain, Vi (=we) is not a possible binder; thus seg selv is free. The only other possibility, coindexing of Ola and seg selv, violates Principle A. Hence, the sentence is ungrammatical. Turning now to ham selv, the embedded clause is the governing category for ham selv, which must be bound to the object. But there is no object in the embedded clause and the coindexing of the NP Ola in the matrix clause with ham selv violates Principle A. Again, the sentence is ruled out. Notice that a pronoun (ham) is allowed in this position, satisfying Principle B. Further, only seg may be in free variation with pronouns (as in (9)); as predicted by the Binding Principles ham selv and seg selv are in complementary distribution with the pronoun ham (as in (6), (7), (8), (10)). The following examples illustrate the above.

ON MARKEDNESS AND BINDING PRINCIPLES

- (6) Ola snakket om seg selv_i/*hamselv_i/*seg_i/*ham_i
Ola talked about (him)self.
- (7) Ola_i lovet meg a snakke om seg selv_i/*ham selv_i/
*seg_i/*ham_i
Ola promised me to talk about (him)self.
- (8) Vi fortalte Ola_i om ham selv_i/*seg selv_i/*seg/
*ham.
We told Ola about himself.
- (9) Ola_i bad oss snakke om seg_i/*seg selv_i/*ham
selv_i/ ham_i
Ola asked us to talk about (him)self/him.
- (10) Ola_i vet at vi snakket om *seg selv_i/*ham
selv_i/*seg_i/ ham_i
Ola knows that we talked about him.

A last example concerns languages where the INFL node does not contain AGR features (cf. Japanese, Korean, Chinese...). In such languages, regardless of whether the sentence is finite or infinitival, a reflexive element in subject position has no accessible SUBJECT in its own clause. Thus, an anaphor may appear in the subject position of a clause and have an antecedent outside the clause in which it appears. The following examples, taken from Dong-Whee Yang(1983) illustrate this case. (i.e. (11) and (12)) Note that the reflexives of these languages (zibun (=self) in Japanese, and caki (=self) in Korean) are bound by any c-commanding subject.

- (11) John_i -wa [Bill_j -ka [Mary_k -no zibun -ni
TOP NM 's self_{i,j,k} DAT
taisuru taito] -o hinansita -to] omotte iru
toward attitude criticize COMP think
(John_i thinks [that Bill_j criticized [Mary's_k
attitude towards self_{i,j,k}]]])

(Japanese)

CELIA JAKUBOWICZ

- (12) John_i -in [Bill_j -i [Mary_k -ka [Tom_l -i'y caki_{i,j,k,l}]
TOP NM NM 's self
-e tehan thæto] -lil silha -n -ta -ko] sengkakha
toward attitude AC gate ASP DEC COMP think
-n -ta -ko] mit -nin -ta
ASP DEC COMP believe ASP DEC
(John_i believes that Bill_j thinks Mary_k hates Tom's_l
attitude towards self_{i,j,k,l})
(Korean)

It is important to note that the antecedent of a reflexive must be a c-commanding subject, unlike the locally bound reflexives and reciprocals. Dong-Whee Yang (1983) relates this fact to the idea that a c-commanding subject is the most prominent NP in the available domain. I will not pursue this issue here.

The data presented above show that the relaxation of the specified subject condition is possible either when the reflexive appears in a clause where the tense element of the INFL node is dependent on that of a higher INFL (regardless of the presence of AGR features as in subjunctive complements of Icelandic and Italian) or when the reflexive appears in a clause where the INFL node does not contain AGR features, regardless of the presence or absence of the [+Past] tense features in the INFL node (as in infinitivals in Norwegian and in languages lacking AGR features in the INFL node (such as Japanese and Korean). Thus languages vary according to whether the AGR features of the INFL node or the type of tense element determine opacity. But crucially, where the AGR element is correlated with a free tense (cf. indicative clauses), the domain is opaque; where Tense and AGR are dissociated, either the tense type or the lack of AGR features in the INFL node license the passing over of a specified subject in the link reflexive-antecedent.

If this analysis is correct, the choices left open by the binding parameter follow from the way in which certain choices are made at the level of the INFL node. It is tempting to suggest, that in the unmarked case, a free tense is one to one correlated with AGR. In this view, the various forms of Tense-AGR dissociation instantiate the values that the INFL parameter can have. Once the INFL parameter is set a cluster of effects follow and have consequences in various subsystems of the grammar. At the level of the binding module with which we are concerned, the consequences show up in the determination of the binding domain. The fact that tense type and AGR (with language specific restrictions) func-

ON MARKEDNESS AND BINDING PRINCIPLES

tion as licensing properties for not locally bound anaphors may be invoked to justify the claim that the properties associated with the binding parameter can be explained, at least partially, in terms of choices made at the level of the INFL node. There is little doubt that the picture sketched here is far from exhaustive. For instance, the licensing properties of Tense type and AGR do not always generalize across the various reflexive NPs that a language might possess (cf. Norwegian and Italian). To account for this fact we might provisionally put the burden on the lexical component of the grammar, assuming that some reflexive NPs are specified so as to be sensitive to the licensing properties. I will not pursue this matter further. On-going research might be able to trace these differences back to some other properties of the grammar; even though it will still be possible to argue that Tense type and AGR are licensing properties for not locally bound anaphors.

Throughout we have assumed that, in the unmarked case, an anaphor is A-bound in its governing category; moreover, that UG leaves open the option of neglecting a specified subject under certain licensing properties, given, as output, instances of non-local binding. On what grounds is this hypothesis justified? In other words, why should an anaphor that is bound in its governing category, rather than one that is not locally bound, constitute the unmarked case?

2. Markedness and Restrictiveness

Linguistic theory, as formulated by Chomsky (1965 and further work) is a biological model for the specific capacity of the human mind to develop a grammar. Thus, one of the primary goals of linguistic theory is to formulate an empirical hypothesis about the structure of the genetic endowment which, it is assumed, constrains children in their hypothesis formation, eliminating many logical possibilities and thereby facilitating the language developmental process. It is a fact that very young children master a rich system of knowledge without reinforcement for what is not a syntactically well-formed sentence (Brown and Hanlon (1970)), without significant instruction (cf. Newport, Gleitman and Gleitman (1977)), and despite the so-called "deficiency of the stimulus" (cf. Wexler and Cullicover (1980), Lightfoot (1982)...). Modern linguistics has to find an answer to this puzzle by assuming that certain properties are available to the organism independently of linguistic experience, which permit language growth to overcome the poverty of the stimulus, and to take place extremely quickly in a surprisingly uniform way. In formulating an hypothesis about the nature of these properties (that is, in the theory of UG) the fact that the child does not have access to negative evidence has been a principal point of concern. Basically,

this fact leads linguists to assume that UG does not make available choices that can only be solved by such evidence. If the hypothesized language were to include the target language, the child would need the information that the excess structures are impossible (that is, those structures that belong to the hypothesized language but not the target language). But, as noted, this evidence is not available.⁷ Thus, linguistic theory must show how, in principle, the child can arrive at the target language on the bases of positive data only. This goal is attained in cases where the hypothesized language is a proper subset of the target language, or when both languages either intersect or are disjoint. In all three cases there is positive evidence available that structures of the target language are not in the hypothesized one (cf. Lasnik (1983)). It is obvious that when the hypothesized language coincides with the target one, no problem arises. Baker (1979, 1981), Dell (1982), Roeper (1981), Berwick (1982) and others have argued that children must hypothesize narrowly; in particular, a properly included language has to be the unmarked case, and hence, the child's default hypothesis.

In a recent computational approach Berwick (1982) made the Subset Property the single source of all the constraints predicted by his theory of syntax acquisition, developed in the framework of Chomsky (1981). Berwick claims that this property may be invoked to justify the setting of parameters and to account for the order in which certain decisions are made during the developmental process. The basic idea underlying the notion of the Subset Property is the following: suppose that a certain parameter of UG has two values: [0] and [1]. Suppose moreover that the choice of [1] gives rise to a set of generated sentences which properly includes the set determined by the value of [0]. Then, by the Subset Property, the child sets the value at [0]; that is, he chooses the narrowest language unless presented with data which show that his initial choice is incorrect.

According to Berwick (p.276), a family of languages possesses the Subset Property when identification in the limit from positive evidence can proceed successfully. Notice, however, if the point of this condition is to ensure that the child's first choice be the smallest language compatible with evidence so far encountered, it follows that the availability of positive data is a necessary condition but not a sufficient one. The Subset Property requires moreover, that the size of the language be specifiable. Berwick argues that the number of output surface strings (of some linguistic level) generated by a grammar (or by a rule of the grammar) specifies the size of a given language. Thus, the Subset Property relies on a rather abstract notion of externalized language, namely the type of output strings

ON MARKEDNESS AND BINDING PRINCIPLES

licensed by certain properties of the program that "writes down" the grammar, under the idealized and simplifying assumption that those properties can be independently considered. This brings us to another important point, namely which criteria apply to specify the number of output strings that form a given language. Berwick's analysis of some cases where the property applies suggests that either the optionality of certain properties, or the scope of a certain rule (or principle), may be taken into account.⁸

Let us now examine how the Subset Property might apply in the case of overt anaphors and pronouns. Recall that the data presented in section 1. permitted us to distinguish: (i) languages in which reflexives are always locally bound (cf. English) and (ii) languages in which reflexives need not be locally bound, depending on the INFL node type of the clause in which they appear (cf. Icelandic, Italian, Norwegian, Japanese..). How does the Subset Property apply? The answer is straightforward.

A grammar that licenses long-distance binding externalizes a broader language than a grammar that does not license it. The former contains two types of output sentences (sentences where anaphors are locally bound and sentences where anaphors are not locally bound). The latter contains only one type of output string (sentences where anaphors are locally bound). Thus, a language for which the grammar only allows local binding is properly included in a language for which the grammar allows the option of non-local binding. According to the Subset Property, a child learning English will have to learn almost nothing, since the hypothesized language coincides with the target one. Children learning Italian or Icelandic will allow the parameter to be set so as to generate a larger language when presented with sentences containing anaphors that are not locally bound.

To sum up: deficiency of the stimulus considerations and the fact that children do not have access to negative data legitimate the claim that a restrictive hypothesis constitutes the unmarked case. Moreover, a particular form of restrictive hypotheses, the Subset Property may be invoked to justify the direction of the Binding parameter setting, namely it is predicted that children exposed to a language in which the unmarked option (cf. local binding) holds will use positive data only to fix the option already present. Children exposed to a language in which the marked option (cf. non-local binding) holds will first assume the unmarked option and only later expand their initial hypothesis, as demanded from positive evidence in their language.

Consider now Berwick's second claim, that the Subset Property may be invoked to justify the order in which the child formulates hypotheses during the developmental process.

Does this property allow the prediction of an order in identifying the properties of referentially dependent NPs such as overt anaphors and pronouns? Notice that an instantaneous acquisition model takes for granted the lexical typological distinction between anaphors and pronouns which permits Binding Principles to apply. The non-instantaneous child has to establish this distinction, that is, he must learn that himself is a "reflexive" expression, that him is a "pronoun", as idiosyncratic facts about English. It is obvious that children learning English neither receive specific instructions nor direct experience about such matters, as they presumably do in the case of names, such as John or table. It is natural, then, to claim that some general syntactic principles provide a basic framework within which, Principles A and B can apply once the lexical distinction is established.

Consider now the binding properties of anaphors and pronouns: Principle A requires an anaphor to be A-bound in its minimal governing category; Principle B requires a pronoun to be A-free in its minimal governing category. Following Berwick we assume that the range of application of Principles A and B permit the binding possibilities of anaphors and pronouns to be ranked in order of decreasing restrictiveness on output surface structures. The strongest constraints are placed on anaphors, since they must be bound in their local governing categories; weaker constraints are placed on pronouns, since they can be bound across arbitrarily large domains. It is clear that the weakest constraints are placed on R-expressions (names), since they may appear anywhere. That is, an anaphor cannot be bound across arbitrary domains, such as in (13).

(13) *They_i knew [that I said [that John shaved themselves_i].

But this is precisely the set of domains where pronouns may appear, as seen in (14).

(14) They_i knew[that I said [that John shaved them_i].]

Thus, pronouns can be linked across large arbitrary domains, whereas anaphors must be bound in a finitely specifiable domain. Notice further, that pronouns may appear in subject and complement position in finite and infinitival embedded clauses, whereas anaphors are excluded from the subject position of a finite clause (unless within NP, which satisfies the well-formedness condition (cf.(3)). One can then see that the set of output type sentences where pronouns may appear is larger than the one where anaphors may appear (in an extended sense, determined by coindexing). According to the Subset Property, it is possible to predict an order in which hypotheses about the properties of referentially dependent NPs will be made. The most restrictive assumption is that an NP is bound in its minimal governing category, hence the default hypothesis should be that an NP is an anaphor. Recall that the child has to establish the meaning of expressions such as "each other" "themselves", "them", and assume that he only knows that they are not names. Suppose now that himself is heard first as in: 10

ON MARKEDNESS AND BINDING PRINCIPLES

(15) John said that Peter washed himself.

By default, himself will be considered to be locally A-bound. Since the child will not be presented with examples where himself is not locally bound, the firstdefault hypothesis will be fixed. Suppose now that the child is presented with (16).

(16) John said that Peter washed him.

If the typological distinction between himself and him is not yet established, the child will assume, by default, that him is locally A-bound. Of course, this hypothesis is wrong. In this case, positive evidence will allow him to correct his initial assumption. The lexical typological distinction established, Principles A and B automatically apply.⁹

To the extent that the above account is correct, the Subset Property authorizes a developmental prediction: if the child's first assumption is that a referentially dependent NP is locally A-bound, then he will not distinguish between anaphors and pronouns, namely he will consider that both expressions are locally A-bound.

Tentatively assuming such a framework I should emphasize here a crucial methodological point. As mentioned above, the Subset Property focuses on a rather abstract notion of externalized language. However, it could hardly be maintained that children learning English (or any other language) possess an actual procedure to compare the set of output types of two languages in order to choose the narrowest one. Thus the Subset Property, even if correct, cannot be assigned the status of an effective procedure. In other words, the Subset Property cannot be a plausible candidate for a real principle of language growth. Rather, this notion can be thought of as a formal computational property that language development may meet. One may ask, for example, what the properties may be of the class of languages that the child masters first. To the extent those properties can be traced back to principles and rules of UG, they might inform us about the unfolding of the system. It seems then reasonable to assume that finding out how language development could work, might prove to have non-trivial consequences for the way we understand how language does work.

Summarizing: the Subset Property, if invoked to justify the Binding parameter setting, permits the prediction that even in languages where the grammar licenses the option

of non-local binding, local binding will nevertheless be the child's first assumption. Cross-linguistic experiments may demonstrate whether this claim is correct or whether it should be discarded. On the other hand, the Subset Property, if invoked to justify a developmental pathway in their interpretation of overt anaphors and pronouns, allows a second prediction, that initially children will assume both types of expressions to be locally A-bound, that is, children will not differentiate the binding properties of one expression from those of the other. To the extent that the second prediction is not disconfirmed by the data, this result might provide indirect evidence in favor of the first. It would be surprising indeed if we were to find that a child interprets both anaphors and pronouns as being locally A-bound in English, while in Italian, for instance, his first assumption is that anaphors are not locally bound.

A variety of experiments, principally concerned with the second prediction have been conducted. Two of them are reported in the next sections.

3. The Experiments

3.1 Experiment I

3.1.1. Experimental Conditions

Four variations in the form of the embedded clause were tested: anaphors (himself, each other) and pronouns (him, them) in object position in finite clauses (as in (17)); anaphors and pronouns in subject position in infinitival clauses (as in (18)); anaphors and pronouns in PPs in finite clauses (as in (19)); and anaphors and pronouns NP-contained in subject position in finite clauses (as in (20)). A fifth set of sentences consisted of simple sentences with anaphors and pronouns within NP in object position as in (21), and anaphors and pronouns in PPs as in (22).

- (17) a. John and Peter said that Sue and Mary washed each other.
 b. John said that Peter washed himself.
 c. John said that Peter washed him.

- (18) a. Peter wanted him/himself/each other/ to kick the ball.¹⁰
 b. John said that Peter wanted him/himself/ each other/to kick the ball.

- (19) John said that Peter put the ball next to him/himself/each other.

- (20) John told Peter that the picture of him/himself/each other/ was on the door.

ON MARKEDNESS AND BINDING PRINCIPLES

- (21) John gave Peter the picture of him/himself/
each other.
- (22) John and Peter put Sue and Mary next to them/
themselves/ each other.

Four different matrix verbs and seven different embedded verbs were used, as seen in (23).

- (23) a. Matrix Verbs: to say, to tell, to give, to
want
b. Embedded Verbs: to wash, to scratch, to tickle
to pinch, to kick, to touch,
to pet, to put

The sentences were presented in four blocks. Four different random orders of presentation were used. The design was completely counterbalanced.

An acting out task was used. The material consisted of 6 articulated puppets which can open their mouths (3 boys and 3 girls), pictures of the puppets, and several toys. Subjects were 7-3, 10-4, and 11-5 year old children who have successfully passed two pre-tests (one of production and one of comprehension), designed to evaluate their knowledge of the subcategorization features of the verbs and syntax of complex clauses. 10 sentences were used in the production pre-test and 14 in the comprehension pre-test. Moreover, each one of the pre-tests and tests was preceded by a context sentence in which a third character was introduced that was not mentioned in the test sentence as seen in (24)

- (24) John, Peter and Dave were playing with mud
and they got dirty. Show me: John said that
Peter washed him.

Before the child began the experiment, the experimenter taught the names of the puppets and the representation of verbs of saying by opening the mouth of the puppet. The subjects responses to the test sentences were scored according to whether the matrix subject was correctly or incorrectly interpreted, and according to whether the antecedent of the pronominal or anaphoric expression was considered bound or free.

3.1.2. Results ¹¹

The main results of this experiment are presented in Figures 1, 2, and 3. An inspection of the data in Figure 1 reveals a striking difference in the pattern of response for pronouns and anaphors for the 3 year old children

CELIA JAKUBOWICZ

(under G.1; G.2 and G.3 correspond to 4 and 5 year olds respectively). Children as young as 3 do not make errors with anaphors. When errors show up they concern pronouns only and in most cases, they consist of binding the pronoun in its governing category as if it were an anaphor. 4 and 5 year old children do correctly interpret anaphors and pronouns in finite clauses. However, when these expressions appear in subject position in infinitival embedded clauses, even 4 and 5 year old children make errors with pronouns, but not with anaphors, as seen in Figure 2. As in the preceding case, the errors consist of interpreting the pronoun as locally bound.

Consider now Figure 3. When presented with ambiguous sentences, with either pronouns or anaphors, the younger the child is, the more likely he is to choose as antecedent the closest c-commanding NP, not only for the anaphors, but also for the pronouns. 5 year old children do differentiate between anaphors and pronouns, namely, [NP,S] is taken as antecedent of himself whereas him and each other are linked either to [NP,S] or to [NP, VP].

Finally, when anaphors and pronouns appear in PP position in finite embedded clauses, children of all ages link not only the anaphor, but also the pronoun to the c-commanding NP in the embedded clause (I have not presented the data for this last case.).

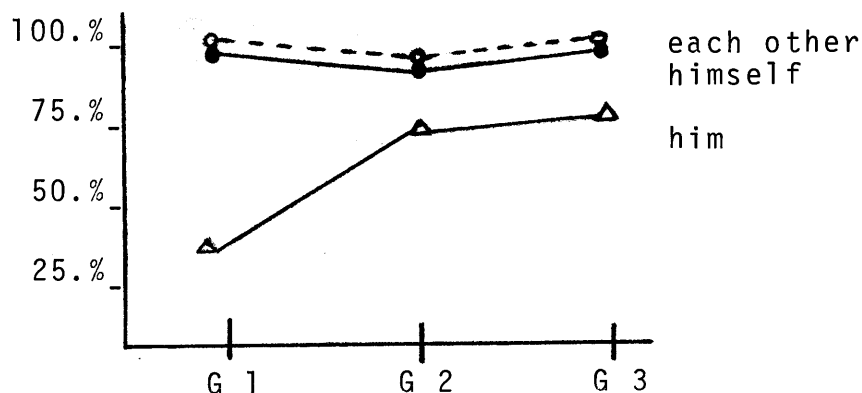


Figure 1

Percentage of correct answers for anaphors and pronouns in object position in finite embedded clauses.

ON MARKEDNESS AND BINDING PRINCIPLES

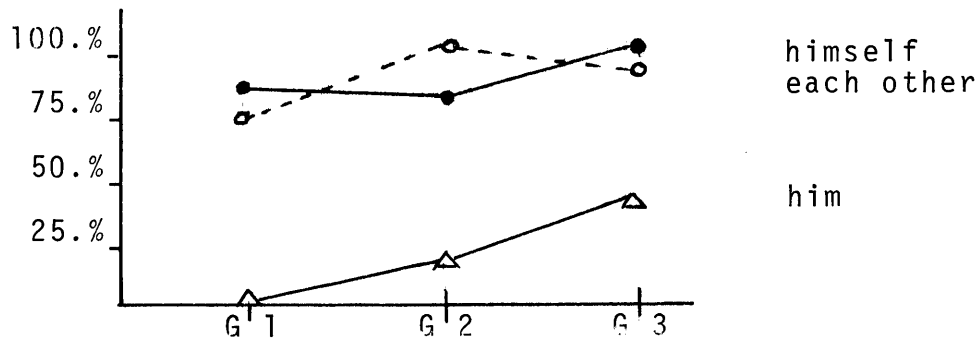


Figure 2.

Percentage of correct answers for anaphors and pronouns in subject position in infinitival embedded clauses.

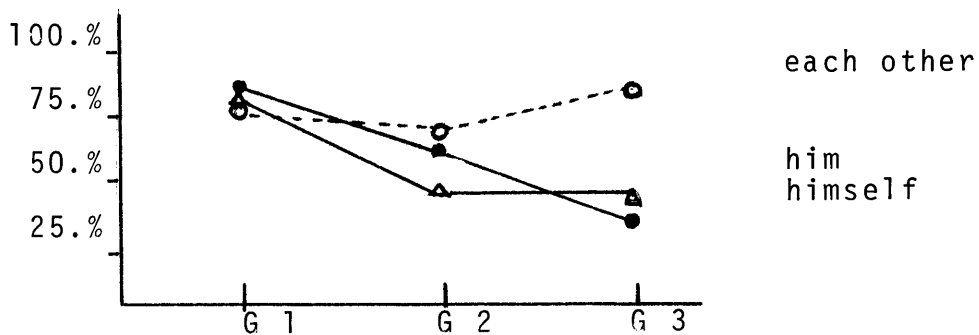


Figure 3

Percentage of answers where anaphors and pronouns are linked to the closest c-commanding NP x Age Group.

- Sentences: (i) John gave Peter the picture of him/himself/each other.
 (ii) John told Peter that the picture of him/himself/each other/was on the table.
 (iii) John and Peter put Sue and Mary next to them/themselves/each other.

To sum up: the results we have obtained confirm the developmental prediction. The fact that 60% of 3 year old children locally bound anaphors as well as pronouns, supplemented by the fact that 75% of these children choose the closest c-commanding NP as antecedent for the three expressions when within NP, can be invoked to justify the claim that young children are constrained by a principle of local A-Binding. Notice however that in the test sentences, dominance and linearity overlap; that is the closest c-commanding NP is, at the same time, the nearest NP to the left of the anaphoric/pronominal expression. Hence, the

CELIA JAKUBOWICZ

results above could be assigned to a linear rather than to a dominance constraint. In order to exclude this alternative interpretation, a second experiment was carried out.

3.2 Experiment II

The object of this experiment is twofold: (i) to distinguish between dominance and linearity and to identify the role played by lexical pronominal features (i.e. gender and number).

Consider the following sentence.

(25) John said that the friend of Dave washed himself.

If in (25) the child obeys a dominance constraint, he will correctly bind the anaphor to the friend, whereas if he obeys linearity, he will incorrectly link the anaphor to Dave.

Consider now the following sentences.

- (26) a. Mary said that John washed her.
 b. John said that Dave and Peter touched him.
 c. Cathy said that Sue and Mary wanted her to kick the balls.

If in the above sentences, the child takes into account the pronominal gender and number features, he will link the pronouns (if interpreted as proximate) to the matrix NP. If the gender/number features are disregarded and the typological distinction between anaphors and pronouns is not yet established, he will incorrectly link the pronoun to the closest c-commanding NP.

The test material consisted of sentences where anaphors and pronouns appear in object position in finite embedded clauses, in object position in infinitival embedded clauses, and in subject position in infinitival embedded clauses. A fourth set of sentences contained pronouns in subject position in finite embedded clauses. The complex NP phrase was presented under two modalities: NP of NP (cf. (25)) and genitival constructions, as in (27).

(27) John said that Dave's friend washed himself.

Moreover, in sentences where anaphors and pronouns appear in subject position, half of the sample was presented with sentences with an overt complementizer, and half with the complementizer deleted (i.e. (28)a,b).

- (28) a. Cathy said that Mary's friend wanted for herself/her to push the cars.
 b. Cathy said that Mary's friend wanted herself/her to push the cars.

ON MARKEDNESS AND BINDING PRINCIPLES

Given that neither the form of the complex NP nor the absence/presence of the for complementizer gave rise to any variation in the children's answers, I do not distinguish these factors in the presentation of the results. Subjects of the experiment were 9- 3, 12- 4, and 10-5 year old children. All other remaining experimental conditions were as in Experiment I.

3.2.2 Results

Figure 4 shows again a striking difference between the comprehension of anaphors and pronouns; namely the correct comprehension of anaphors is well established in 3 year old children, the correct interpretation for pronouns is only reached by older children.

Consider now Figure 5. It shows children do not make errors when presented with sentences containing anaphors, neither in the case of infinitivals nor in the case of finite embedded clauses. That is, children do correctly link the anaphor to the c-commanding NP in its governing category and not to the NP which is minimally distant from the expression in the linear sequence. Turning now to Figure 6, it shows that errors concern pronouns only, and, as in the preceding experiment, they consist of linking the pronoun in its governing category to a c-commanding NP. Further, when pronouns appear in subject position in infinitival embedded clauses, these errors are very frequent and observed even in older children. However, as seen in Figures 7 and 8, when the nominal element of the lower clause does not have the same gender or number features as the pronoun, only younger children persist in considering the pronoun as not free (note that number, but not gender is concerned here). Moreover, the comparison between Figures 7 and 8 shows that the (facilitatory) effect of the agreement features interacts with the tense of the embedded clause. Finally, the data in figures 9 and 10 give further support to the hypothesis that when children know neither the typological properties of an element, nor the agreement features, they interpret these elements as being locally A-bound.

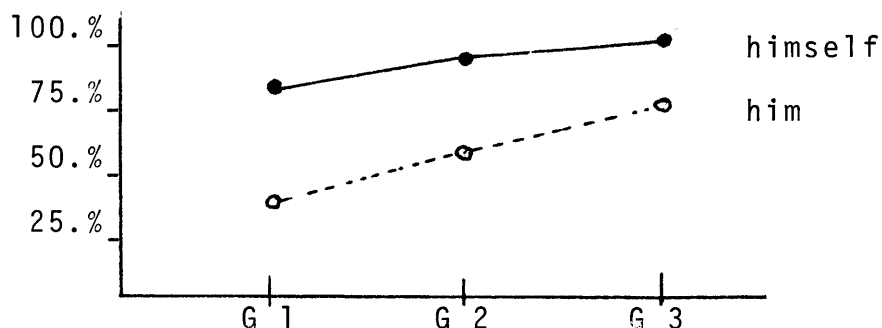


Figure 4 Percentage of correct answers for anaphors and pronouns x Age Group

CELIA JAKUBOWICZ

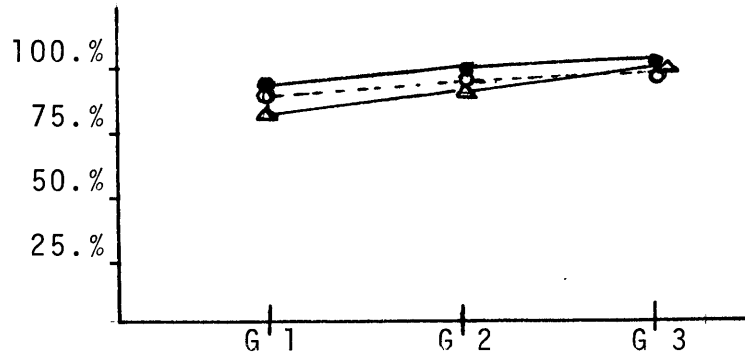


Figure 5 Anaphors

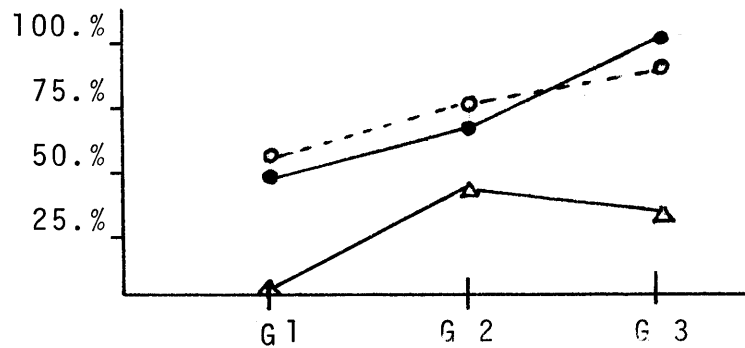


Figure 6 Pronouns

Figures 5 & 6

Percentage of correct answers for anaphors and pronouns x Age Group, taking into account the syntactic position of the expression and the inflection (finite or infinitive) of the embedded clause.

- ——— ● Object position, finite embedded clause
- - - - ○ Object position, infinitival embedded clause
- △ ——— △ Subject position, infinitival embedded clause

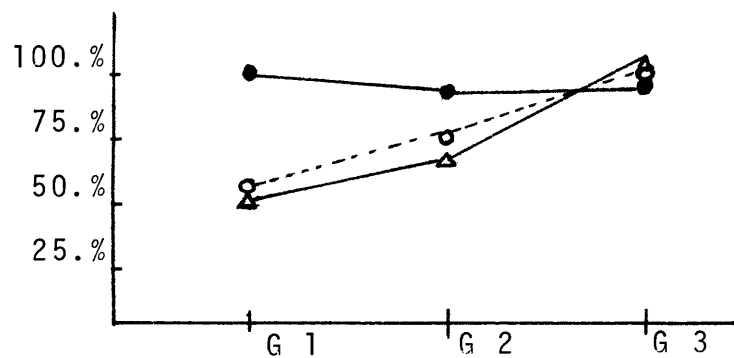


Figure 7

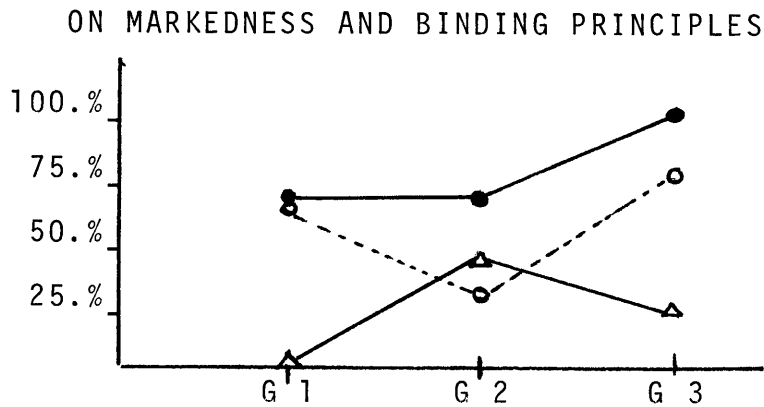


Figure 8

Figures 7 & 8

Percentage of correct answers for pronouns in finite embedded clauses (Figure 7) and in infinitival embedded clauses (Figure 8), x Age Group x Gender and Number.

Figure 7

Sentences:

- ▲ ——— ▲ John said that Dave washed him.
- ——— ● Mary said that John washed her.
- - - - - ○ John said that Dave and Peter touched him.

Figure 8

Sentences:

- ▲ ——— ▲ Peter said that John wanted him to push the car.
- ——— ● John said that Cathy wanted him to pet the cat.
- - - - - ○ Cathy said that Sue and Mary wanted her to kick the balls.

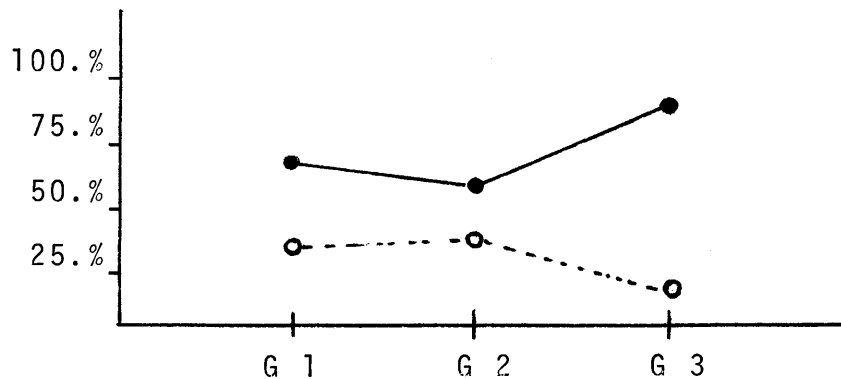


Figure 9

- ——— ● Pronouns linked to matrix subject
- - - - - ○ Pronouns linked to matrix object

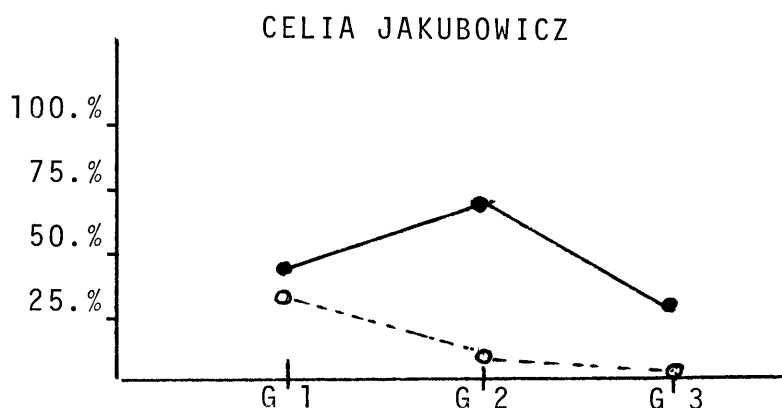


Figure 10

●——● Linked to matrix subject and object
○-----○ Linked to matrix object

Figures 9 & 10

Choice of antecedent for pronouns in finite embedded clauses

Figure 9 Dave told John and Peter that he will kick the ball.

Figure 10 John told Peter that they will pet the cat.

To summarize: very young children do not distinguish between anaphors and pronouns from a typological point of view. But formal operations, for which they seem to have very little choice, put conditions on how referentially dependent expressions are interpreted. That is, linking through c-command in a local domain is present from the beginning of the process; in an English-like language, given that no example where each other or himself is not locally bound ever occurs, the default hypothesis is retained, whereas for pronouns, positive evidence shows that the default hypothesis is incorrect. Thus the Binding Principles automatically apply once the typological distinction of the expressions considered here is established, triggered by positive evidence. In other words, knowledge of pronominal features (gender, number) is not sufficient to determine the Binding properties of pronouns and anaphors, nor is the knowledge of the syntax of embedded clauses.¹² Furthermore, the above data lend support to the hypothesis that verbs of the "want" type do not permit S' deletion and that, in the unmarked case, the subject of an infinitival is PRO. Notice that a language that permits S' deletion is broader than a language that does not. According to the Subset Property, the default assumption should be that S' deletion is not possible. This prediction is supported by our data.

4. Conclusion

The work I have reported provides strong evidence in support of the hypothesis that a principle of local A-binding constrains young children's interpretations of referentially dependent NPs. Then, it permits us to argue against current

ON MARKEDNESS AND BINDING PRINCIPLES

psycholinguistic hypotheses according to which the first linguistic principles are linear rather than structural. Further, the fact that similar restrictions of dominance and locality have been attested in children's interpretations of sentences involving other types of expressions and syntactic processes (cf. control, coordination, backward pronominalization)¹³ legitimize the claim that this principle forms part of the schematism that constrains language growth. Consequently, it provides new evidence in support of the theory of UG. Note, however, that such a claim could be denied (as has often been the case in the developmental literature). The fact that young children locally bind anaphors as well as pronouns, could be taken as evidence in support of the hypothesis that children do not obey Binding Principles, from which it could follow that either Binding Principles are wrong, or that they are not part of the genetic endowment or more strongly, that no innate system exists at all.

It is worthwhile to note that such a conclusion arises from a deep misunderstanding of the significance of the theory of UG. Since the theory conceives of language development as an instantaneous process, it takes for granted, in the formulation of the principles, the specification of data which, it is obvious, relies on linguistic experience. In the case of the Binding Principles, for example, the model assumes the typological distinction between anaphors and pronouns which, we have seen, is not yet available for young children. Since such a distinction is attained through experience, it could hardly be maintained that children's behavior provides direct evidence against or in support for UG.

Consider again the Binding Principles: they are well-formedness conditions in which a certain property, as well as the element which has to meet the property, are specified. The property could be regarded as a statement schema, containing exactly one unbound variable as in (29)a,b.

- (29) a. x is A-bound in its governing category.
 b. x is A-free in its governing category.

The well-formedness condition would be satisfied when a constant which replaces x in (29)a & b is a true instantiation of the property.

Consider now the child. If Binding Principles form part of UG, he is equipped with schemata (29)a&b. But in order to replace the variables by the appropriate constants, the child needs to identify from among the elements of the language to which he is exposed, those which can be inserted into the variable slot. Our results permit us to assume

CELIA JAKUBOWICZ

that the child knows that expressions such as himself, each other, him, are not R-expressions (cf. they are not names of trees!) However, this is not enough. He needs to "learn" that himself is a reflexive, that each other is a reciprocal, and that him is a pronoun, in order to insert them as constants in the appropriate variable slot. This task requires time and a certain density of experience. From the fact that the child does locally bind him and himself, the only conclusion that can be drawn is that when he does so, he does not yet honor the relevant typological distinction between anaphors and pronouns. In other words, he has not yet identified the constants with which to replace the variable in the schemata above. But, as we have seen, five year old children (and also some 3 and 4 year olds) locally bind himself and link him to an antecedent outside the local domain. That is, once the distinction is established, Binding Principles automatically apply. From there it is possible to conclude that our results provide indirect evidence in support of UG. The evidence is indirect since we are imputing existence to the schemata based on the fact that Binding Principles become operative once the constants are identified. Note, if such schemata did not exist at all, it would be surprising indeed that children were able to arrive at the relevant typological distinction. Hence, linguistic experience is a necessary but not sufficient condition for performing such a task. We conclude then, that the child is equipped with such schemata and that a principle of local A-binding provides a basic framework within which the lexical distinction is established. In other words, this principle applies to the data in such a way that, upon further positive evidence only, the identification of himself as an anaphor and that of him as a pronoun is set up. In the case of himself, positive evidence shows that the outcome of the principle of local A-binding is correct. The initial assumption is then fixed. In the case of him positive evidence shows that locally binding this expression is wrong. The initial assumption is then corrected. In other words, first, referentially dependent expressions (and presumably empty categories) are locally bound. If we were to refer to the typology of categories assumed in UG (Chomsky (1982)), we could argue that first, the typology is limited to R-expressions and non-R-expressions. To rephrase our claim, positive evidence, through the operation of the principle of local A-binding, fixes the principle which requires that certain expressions of the second type (anaphors) must be locally bound, and triggers the principle which specifies the domain where certain other expressions of the second type (pronouns) must be A-free. Moreover, positive evidence relaxes the domain where certain other expressions of the second type (long-distance anaphors) may be not locally bound. In this view, the instantaneous model is consistent with the non-instantaneous child. Under the assumption that the child is equipped

ON MARKEDNESS AND BINDING PRINCIPLES

with the two schemata, the application of the principle of local A-binding to linguistic experience, permits an explanation of the attested developmental sequence. No other factors (cf. semantic or pragmatic) are needed to interpret how, in this case, language grows.

Let us now briefly consider one last issue: does the empirically attested pathway reflect the structure of the linguistic theory? As for the setting of parameters, there is reason to believe that an affirmative answer could be given to the above question if we were to confirm, by means of cross-linguistic experiments, the developmental prediction that our results suggest we should expect. In other words, the order in which children consider the parametric choices would reflect the implicational structure of the theory. To mention one example: suppose that Icelandic children first interpret sig as being locally A-bound only, and upon further positive evidence they admit sig to be also non-locally bound. In this case, the developmental sequence could be regarded as a reflection of the theory which assumes that a theory of the steady state (cf. the theory of grammar) is derived by fixing, through experience, the values of parameters left open in the theory of the initial state (cf. the theory of UG).

As for language growth, the question is more complex. The fact that the full operation of Binding Principles is contingent on the earlier operation of a principle of local A-binding, could be related to a maturational change in another system, like the size of the working memory. If it were so, the attested developmental pathway could not be considered a direct reflection of the structure of the model. This is an open question which deserves much more discussion.

Finally, as for the Subset Property, which we have adopted to establish our prediction, the most we can provisionally claim is that it has not been discarded by the data. As mentioned before, the property itself cannot be regarded as an actual procedure of language growth (but the principle of local A-binding can.) Notice, however, that if the predictions allowed by the Subset Property are empirically attested, the developmental sequence becomes characterizable in terms of the rules and principles proposed by the linguistic theory. That is the reason why it is appealing to adopt such a framework in developmental research. Again, it is an open question whether or not the predictions that this property licenses with respect to other modules of the theory are true. Further research might provide evidence to answer such questions.

CELIA JAKUBOWICZ

FOOTNOTES

*This research was partly sponsored by a fellowship of the Fondation Fyssen (Paris, France) and by the MIT Center for Cognitive Science. I would like to thank A. Honda and R. Mount for their assistance in the collection of data, and the principals, teachers and children from the MIT Day Care Center (Cambridge MA) and the Technology Children's Center (Cambridge MA) for their invaluable collaboration. Special thanks should go to N. Chomsky for discussing and improving each step of this research. Thanks also to J. Aoun, A. Belletti, L. Burzio, M. Halle, J. Maling, A. Weinberg and M.L. Zubizarreta for encouragement and criticism, to B. Levin for checking my English, and to D. Massam for typewriting this paper.

¹Throughout I am concerned with sentences in which, either an anaphor or a pronoun cooccur with one or more definite noun phrases, and the question is whether or not it is possible to establish a referential link between them. The term "referentially dependent" is used to allude to the fact that "proximate" pronouns and anaphors share the property of having a linguistic antecedent.

²I refer the reader to Chomsky (1981, 1982) and references quoted therein for relevant discussion concerning how the Binding Principles work, and the problems for them (eg. pronouns within NPs) which the current formulation leaves unsolved. See also Bouchard (1982), Huang (1982), and Lebeaux (1983).

³See Dong-Whee Yang (1983)

⁴I should emphasize that I disregard some language specific restrictions, with the leading idea being to trace the attested differences back to a unifying principle (See references.)

⁵See Anderson (1982), Johnson (1983), Maling (1981), and Thrainson (1976).

⁶Note that in Italian the emphatic reflexive pronoun "se stesso" (=himself) must be locally bound, regardless of tense type. The same holds for the clitic "si" (=him/herself) (cf. Giorgi 1983)

⁷Besides the psycholinguistic evidence mentioned in the text, results from learnability theory show that admitting negative evidence broadens the class of "possible learnable languages" so much that one can learn languages which cannot be characterized as human because of their rule systems. (Gold (1967))

⁸The question of the criteria which apply to determine the size of a language deserves a more extensive discussion which I will not pursue here. Consider one example where the optionality of a rule in the grammar makes a language broader: Spanish versus English with respect to the null subject parameter.

ON MARKEDNESS AND BINDING PRINCIPLES

FOOTNOTES CONT'D

Spanish is broader than English since in Spanish, but not in English, lexical subjects in finite sentences can be missing. Note that under Rizzi's (1982) account, in a null subject language the AGR element is optionally specified: [+pronoun]. To illustrate the second criterion, scope of application specified in a rule, consider the following case: identification of empty categories such as trace and PRO. Note first that "scope of application" refers to the "space" or "domain" specified in the rule (or principle). If it is strictly restricted, the rule will output only one type of surface string. If it is not, the number of surface strings will be larger. Consider now the case of an empty category such as trace: being an anaphor, the trace must be bound in a local domain. Hence, the number of output type strings will be just one. Consider now the empty category PRO. Since PRO does not have a minimal governing category, the number of output type strings in which PRO is co-indexed with some other element is much larger.

⁹It is a question of fact how much children learn from "indirect negative evidence"; that is, from the non-appearance of sentences that they might expect would appear if their assumptions were correct. Notice that in the above cases, given that anaphors and pronouns are in nearly complementary distribution, the child should be able to start with the assumption that himself and him are not locally linked, and then use a principle of indirect negative evidence to correct the overgeneralization for anaphors. In other words, the non-appearance of forms such as (i)

(i) John knows that Mary likes himself.

would be sufficient to inform the child that non-local co-indexing for anaphors is incorrect. Throughout I assume that the child relies on positive evidence only. It seems to me very difficult to experimentally investigate how indirect negative evidence operates (See Berwick (1982) and Lasnik (1983)).

¹⁰In sentences containing each other, the noun phrases are always plural, as in (17)a.

¹¹Throughout this paper statistical significance is based on the Cochran Q test for related samples, at the .05 level of significance or less.

¹²Matthei (1981) and Otsu (1981) examined children's interpretations of anaphors and pronouns. Matthei's experiments concern reciprocals. He observed that in 64% of the cases, children incorrectly interpret sentences such as (i). (The error consisted in having the horses and the cows jumping over each other).

(i) The horses said that the cows jumped over each other.

However, children did correctly interpret sentences as (ii)

(ii) The boys jumped the cows into each other.

CELIA JAKUBOWICZ

FOOTNOTES CONT'D

As mentioned also in Otsu (1981) and Solan (1983), Matthei's results are not conclusive since he did not experimentally control for the child's comprehension of the syntax of complex clauses.

Otsu (1981) shows that children who understand the syntax of complex clauses correctly interpret anaphors and pronouns. He concludes that once the child masters the syntax of complex clauses, Binding Principles automatically apply. However, a more careful analysis of Otsu's results shows that children made fewer errors involving binding in sentences containing anaphors than in those containing pronouns. Further, it is worth noting that, unlike Matthei's and Otsu's studies, our study does not require that children make adult interpretations of the test sentences in order to demonstrate the existence of genetic constraints. Rather, the nature and distribution of the children's errors justify the claim that innate principles constrain language growth.

¹³See Hsu, Cairns and Fiengo (1983) for a linearity based principle. See Goodluck and Tavakolian (1982), Cohen Sherman (1983), Lust & Mangione (1983), and references cited in these for evidence in support of dominance constraints.

ON MARKEDNESS AND BINDING PRINCIPLES

REFERENCES

- Anderson, S. (1982) "Types of Dependency in Anaphors: Icelandic (and other) Reflexives" Paper presented at GLOW: Paris, France.
- Baker, C.L. (1979) "Syntactic Theory and the Projection Problem" Linguistic Inquiry 10:4. MIT Press.
- Berwick, R.C. (1982) Locality Principles and the Acquisition of Syntactic Knowledge MIT Phd. Dissertation.
- Bouchard, D. (1982) On the Content of Empty Categories MIT Phd. Dissertation.
- Brown, R. and C. Hanlon (1970) "Derivational Complexity and Order of Acquisition in Child Speech" in J.R. Hayes, (ed.) Cognition and the Development of Language John Wiley, N.Y.
- Chomsky, N. (1965) Aspects of the Theory of Syntax MIT Press, Cambridge MA.
- (1975) Reflections on Language Pantheon Books, N.Y.
- (1981) Lectures on Government and Binding Foris Publications, Dordrecht.
- (1982) Some Concepts and Consequences of the Theory of Government and Binding MIT Press, Cambridge, MA.
- Cohen Sherman J. (1983) The Acquisition of Control in Complement Sentences: the Role of Structural and Lexical Factors Cornell University Phd. Dissertation.
- Dell, F. (1981) "On the Learnability of Optional Phonological Rules" Linguistic Inquiry 12 MIT Press.
- Dong-Whee Yang (1983) "The Extended Binding Theory of Anaphors" Paper presented at GLOW: York, England.
- Giorgi, A. (1983) "Toward a Theory of Long Distance Anaphors: a GB Approach" ms, Consiglio Nazionale delle Ricerche, Roma.
- Gold, E. (1967) "Language Identification in the Limit" Information and Control 10.
- Goodluck, H. and L. Solan, eds. Papers in the Structure and Development of Child Language Acquisition University of Massachusetts Occasional Papers in Linguistics, 4. Amherst, MA.

CELIA JAKUBOWICZ

REFERENCES CONT'D

- Goodluck, H. and S. Tavakolian (1982) "Competence and Processing in Children's Grammar of Relative Clauses" Cognition 11.
- Hellan, L. (1980) "On Anaphora in Norwegian" ms, U. of Trondheim/ U. of Massachusetts, Amherst.
- Hsu, J, H. Cairns, and R. Fiengo (1983) "The Development of Grammars Underlying Children's Interpretation of Complex Sentences" ms, The William Paterson College of New Jersey, Queens College, and The Graduate Center of the City University of New York.
- Huang, C.T. (1982) Logical Relations in Chinese and the Theory of Grammar MIT Phd. Dissertation Cambridge, MA.
- Johnson, K. (1983) "Some Notes on Binding in Icelandic and Subjunctives" ms, MIT, Cambridge, MA.
- Lasnik, H. (1983) "On Certain Substitutes for Negative Evidence" ms, U. of Connecticut.
- Lebeaux, D. (1983) "Locality and Anaphoric Binding" ms, U. of Massachusetts, Amherst.
- Lightfoot, D. (1982) The Language Lottery: Towards a Biology of Grammars MIT Press, Cambridge, MA.
- Lust, B. and I. Mangione (1983) "The Principal Branching Direction Parameter in First Language Acquisition of Anaphora" Proceedings of ALNE/NELS 13.
- Maling, J. (1981) "Non-Clause-Bounded Reflexives in Icelandic" Paper presented at NELS 6.
- Matthei, E. (1981) "Children's Interpretation of Sentences Containing Reciprocals" in S. Tavakolian (ed) Language Acquisition and Linguistic Theory MIT Press, Cambridge, MA.
- Newport, E., H. Gleitman, and L. Gleitman. (1977) "Mother, I'd Rather Do It Myself: Some Effects and Non-effects of Maternal Speech Style" in C. Snow and C. Ferguson (eds.) Talking to Children: Language Input and Acquisition Cambridge University Press, N.Y.
- Otsu, Y. (1981) Universal Grammar and Syntactic Development in Children: Towards a Theory of Syntactic Development MIT Phd. Dissertation.

ON MARKEDNESS AND BINDING PRINCIPLES

- Roeper, T. (1981) "On the Deductive Model and the Acquisition of Productive Morphology" in C. Baker and J. McCarthy, (eds) The Logical Problem of Language Acquisition MIT Press, Cambridge, MA
- Rizzi, L. (1982) "Negation, WH-Movement, and the Null Subject Parameter" Foris Publications, Dordrecht.
- Solan, L. (1983) Pronominal Reference, Child Language and the Theory of Grammar D. Reidel Publishing Co. Dordrecht.
- Thrainsson, H. (1976) "Reflexives and Subjunctives in Icelandic" Paper presented at NELS 6.
- Wexler, K. and P. Culicover (1980) Formal Principles of Language Acquisition MIT Press, Cambridge, MA