



1984

## Clitic doubling and M-chains in Pirahã

Daniel L. Everett  
*SIL-UND*

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: <https://commons.und.edu/sil-work-papers>



Part of the [Linguistics Commons](#)

---

### Recommended Citation

Everett, Daniel L. (1984) "Clitic doubling and M-chains in Pirahã," *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session*: Vol. 28, Article 2.

DOI: 10.31356/silwp.vol28.02

Available at: <https://commons.und.edu/sil-work-papers/vol28/iss1/2>

This Article is brought to you for free and open access by UND Scholarly Commons. It has been accepted for inclusion in Work Papers of the Summer Institute of Linguistics, University of North Dakota Session by an authorized editor of UND Scholarly Commons. For more information, please contact [und.common@library.und.edu](mailto:und.common@library.und.edu).

# CLITIC DOUBLING AND M-CHAINS IN PIRAHĀ\*

by

Dan Everett

- 1 Introduction
  - 2 A sketch of Pirahã surface syntax
    - 2.1 Verb morphology
    - 2.2 Phrase structure
    - 2.3 Intrasentential reference
  - 3 The theoretical framework
  - 4 Towards an analysis of intrasentential reference in Pirahã
    - 4.1 A failed analysis
    - 4.2 An analysis of clitic doubling in Pirahã
      - 4.2.1 A note on the nature of clitics
      - 4.2.2 Clitics, ECs and M-chains in Pirahã.
- Appendix: Clitics, pronouns and affixes

## 1 Introduction

Recent studies of syntax have shown clitics to be a rich source of insights into a variety of principles governing the well-formedness of sentences (cf., *inter alia*, Kayne (1975); Steele, et al. (1981); Jaeggli (1982); Borer (1981); Kaisse (1982)). One particularly interesting focus of discussion continues to be the relationship between pronominal clitics and coreferential NPs in the same clause, so called **clitic doubling** constructions, as in (1) (from River Plate Spanish, Chomsky (1981:277)):

- (1)       $Lo_i$  vimos a  $Juan_i$  <sup>1</sup>  
          $him$  we:saw to John  
         We saw John.

At least two points of consensus have been reached by studies on clitic doubling in the **Government and Binding Theory** (henceforth GBT) of Chomsky (1981; 1982). These are : (i) the basic properties of this phenomenon are derivable from the **subtheories** of **Case**, **government**, **thematic (θ) roles**, and **binding** and, closely related to this, (ii) clitic doubling is possible only when a special **Case assigner** appears to attribute Case to the **doubled NP**, since clitics "absorb" Case.<sup>2</sup>

The present paper may be seen as a contribution to this general discussion to the degree that we are able to establish that while the basic properties of clitic doubling in Pirahã are indeed derivable from subtheories of GBT, additional principles are required to explain why languages such as Pirahã can allow clitic doubling without a special Case assigner. The basic facts to be dealt with in this

regard are found in (2)-(4) below. In (2) we see for example that Pirahã pronominals are not distinguished morphologically between reflexives and nonreflexives or, alternatively, that they are free in reference when no doubled NP is present. (3) and (4) show that these pronominals may cooccur with subject and/or object NPs in which case they are obligatorily coreferential with the doubled NP. See the next section for additional data.

- (2) a.  $hi_{3(i)} hi_{3(i/j)} xib -\acute{a}o -b -\acute{a}$   
 $3(i) \quad 3(i/j) \quad hit \quad -telic \quad -perf \quad -rem$   
 (i) 'He hit himself.' :u3  
 (ii) 'He hit him.'
- b.  $ti_{1(i)} ti_{1(i)} xib\acute{a}ob\acute{a}$   
 $1(i) \quad 1(i) \quad hit$   
 'I hit myself.'
- c.  $g\acute{i}_{2(i)} g\acute{i}_{2(i)} xib\acute{a}ob\acute{a}$   
 $2(i) \quad 2(i) \quad hit$   
 'You hit yourself.'

#### Pseudo-topicalization

- (3) a.  $kohoibi\acute{i}hai_{(i)} hi_{3(i)} ti_{1(j)} xib\acute{a}ob\acute{a}$   
 $kohoibi\acute{i}hai_{(i)} \quad 3(i) \quad 1(j) \quad hit$   
 'Kohoibiíhai hit me.'
- b.  $kohoibi\acute{i}hai_{(i)} hi_{3(i)} g\acute{i}_{2(j)} xib\acute{a}ob\acute{a}$   
 $kohoibi\acute{i}hai_{(i)} \quad 3(i) \quad 2(j) \quad hit$   
 'Kohoibiíhai hit you.'
- c.  $kohoibi\acute{i}hai_{(i)} hi_{3(i)} hi_{3(i/j)} xib\acute{a}ob\acute{a}$   
 $kohoibi\acute{i}hai_{(i)} \quad 3(i) \quad 3(i/j) \quad hit$   
 (i) 'Kohoibiíhai hit himself.'  
 (ii) 'Kohoibiíhai hit him.'

#### Complex reference

- (4) a.  $kohoibi\acute{i}hai_{(i)} hi_{3(i)} xabagi_{(j)} hi_{3(j)}$   
 $koibi\acute{i}hai \quad 3(i) \quad xabagi \quad 3(j)$   
  
 $xib\acute{a}ob\acute{a}$   
 $hit$   
 'Kohoibiíhai hit Xabagi.'
- b.  $kohoibi\acute{i}hai_{(i)} xabagi_{(j)} hi_{3(i)} hi_{3(j)}$   
 $Kohoihiíhai \quad Xabagi \quad 3(i) \quad 3(j)$   
  
 $xib\acute{a}ob\acute{a}$   
 $hit$   
 'Kohoihiíhai hit Xabagi.'

We will have more to say about these examples in subsequent sections. Further, we will suggest that the analysis of clitic configurations in Pirahã offers important new insights into parameters governing clitic configurations in **Universal Grammar** (UG).

The discussion is organized as follows. First, a basic sketch of Pirahã surface syntax is provided, focussing especially on intrasentential reference configurations such as those exemplified by (2)-(4) above. This is followed by a brief introduction to the relevant principles of GBT. Next, we propose an analysis of clitic configurations in Pirahã based on the notion of **M-chains**, by which, it is argued, clitics in languages like Pirahã transmit Case,  $\theta$  roles, and other features to their doubled NPs. In the last section the predictions of this idea are tested with regard to Pirahã and shown to be superior to those of other recent analyses of clitics. Finally, an appendix is added in which we speculate on the implications of this analysis for the understanding of how clitics relate to the development of verbal affixes.

## 2 A sketch of Pirahã surface syntax

### 2.1 Verb morphology

It is worth noting here that Pirahã verbs lack two basic features relatively common to verbs in general, namely, they are not marked for either tense or agreement. Thus, in example (2) above, the verb form is constant for 1, 2, or 3 person. Consider as further evidence of this, (5) and (6) below.

- (5) a. xipóihí baóhoipaí koho -ái -p -i -háí  
       woman ocelot eat -atelic -imp -prox -relcert  
       'The woman is/will shortly be/just was eating the ocelot.'
- b. baóhoipaí xipóihí kohoáipiháí  
       'The ocelot is/will shortly be/just was eating the woman.'
- (6) a. hi soxóá kahá -p -i -í  
       3 already go -imp -prox -compcert  
       'He already left/is going.'
- b. ti soxóá kahápií  
       1 already go-imp-prox-compcert  
       'I already left/am going.'

The temporal ambiguity or vagueness of such examples is resolved mainly through context (although some aspectual combinations such as **telic + perfective** force a particular (in this case past)

interpretation. Time words do not offer much help since they are equally ambiguous, e.g. **xahoapió** 'another day' (used either as 'yesterday', 'tomorrow', or 'some other day') and **soxógiái** 'big time' (i.e. 'a long time ago' or 'a long time from now in the future'). We have more to say on agreement in Pirahã in Sect. 4, suggesting that clitics fulfill this function.

Pirahã verbs are richly inflected for aspectual distinctions. In Everett (to appear), some sixteen classes of verbal suffixes are listed, with over thirty distinct aspects and moods. However, we will not go into this system here, since it has little bearing on the present analysis (although, as is mentioned in subsequent discussion, perhaps aspect is best analyzed as generated directly on the verb, rather than in what GBT normally considers to be the position of inflection).

## 2.2 Phrase structure

In Everett (to appear), we argue that basic word order in Pirahã is SOV. This conclusion is based primarily on (i) frequency of this order in relation to other observed configurations and (ii) the importance of word order to the interpretation of grammatical relations. It is also noted that phrase structure in Pirahã shows features common to OV languages, such as postpositions and genitive-head noun orders. To illustrate the relation between word order and grammatical relations, consider the following examples:<sup>4</sup>

- (7) a. kohoibiíhai kapi xiti -bai  
 kohoibiíhai coffee drink -int  
 'Kohoibiíhai drinks a lot of coffee.'
- b.\* kapi kohoibiíhai xitibai  
 \* 'Coffee drinks a lot of kohoibiíhai.'
- (8) a. baíxi xigagaisi xogió kai -p -á -há  
 father God all make -imp -rem -compcert  
 'Father God made everything.'
- b.\* xogió baíxi xigagaisi kaipáhá  
 \* 'everything made God'

Implications of other deviations from SOV order are discussed later in this section and Sect. 4. (9) and (10) illustrate intransitive clauses.

- (9) páxaihi xisáxoi -bai  
 rooster crow -int  
 'The rooster crows a lot.'

- (10) xogihíai kaob -í  
 star fall -prox  
 'The star falls.'

Although a large number of other constituents and structures might be exemplified, here we mention only the set of categories which occur in the syntactic **oblique** position. This node is, in linear terms, the third position rightward in the sentence, following subject and subject clitic positions. Alternatively, it is the first node dominated by VP. Elements occurring in this position (except for proper nouns and a few other isolated cases) are marked by the suffix -o 'oblique' (cf. Everett (to appear) for more argumentation and exemplification). Elements occurring in this position include indirect objects, postpositional phrases, comitatives, and adverbial expressions as in (12)-(16), respectively,

- (12) xoi tabó ap-ó xítixisi xihí-a-há  
 xoi board head(pre)-obl fish put-rem-compcert  
 'Xoi put the fish on top of the table.'
- (13) kaioá xahaigí xigí -o xopí -itar -há  
 kaioá brother with -obl go -iter -compcert  
 'Kaioá left with (his) brother.'
- (14) hoagaixóxai pí -o kahá -p -í  
 hoagaixóxai also -obl go -imp -prox  
 'Hoagaixóxai left also.'
- (15) hiaitíhí xahoigí -o xísa -xop -í  
 Pirahã evening -obl sing -go -prox  
 '(the) Pirahã go sing (in) (the) evening'
- (16) kagáihíai báhiigí baái koab -ái -p -í  
 jaguar slowly wild pig kill -atelic -imp -prox  
 'The jaguar slowly kills the wild pig.'

One further point of interest for the present discussion is that Pirahã verbs may appear without their full complement of arguments, as in (16) and (17):

- (17) speaker A: hi kao xítiixisi kaoap -áp -í  
 3 inter fish hook -punc -prox  
 'Did he catch any fish?'
- speaker B: soxóá kaoapápi  
 already  
 '(He) already caught (some fish).'

- (18) kahá -p -i -hiab -iig -á  
 go -imp -ep -neg -cont -rem  
 '(He) hasn't gone yet.'

This fact about Pirahã sentence structure will be of relevance to the discussion in the fourth section. The property of allowing fewer (overt) arguments than called for by the lexicon is known as the **pro-drop** parameter.<sup>5</sup> More on this is given in subsequent discussion.

The basic structure of noun phrases is: (Genitive) - Head - (Modifier) - (Determiner), as in (19) and (20).

- (19) baíxi kaiíi  
 parent house  
 'father's/mother's house'
- (20) xoogiái sitái biísai gáihi  
 xoogiái feather red that  
 'That red feather of Xoogiái.'

Thus, we may sum up our brief survey of Pirahã phrase structure by the following rules:

- (21) a.  $S \rightarrow N'''' -INFL - V''$   
 b.  $V'' \rightarrow \left\{ \begin{array}{l} N'''' \\ P'''' \\ \text{Particle} \\ \text{Adv} \end{array} \right\} -V''$   
 c.  $V'' \rightarrow (N''''') -V$   
 d.  $N'''' \rightarrow N'''' - (\text{Det})$   
 e.  $N'''' \rightarrow N'''' - (\text{Mod})$   
 f.  $N'''' \rightarrow (N''''') - N$

### 2.3 Intrasentential reference

To understand reference in Pirahã, we must first say a few words about the pronoun system. As will become clear in the course of this paper, Pirahã pronominals include both **pronouns** in the usual sense of the word, and **clitics** (cf. Kayne (1975), Borer (1981) and Sect. 4.2.1 below). Since the notion of clitics relies, however, on the analysis of these elements in GBT, we will therefore continue to use the more neutral term **pronominal**, here.

Let us consider first the simplicity of this system:<sup>6</sup>

- (22) a. ti 'first person singular'  
 b. gi(xai) 'second person singular'  
 c. hi(apióxiái) 'third person singular/non-definite'  
 d. góí 'second person singular imperative'  
 e. kaxáo 'first person plural hortatory'  
 f. xogiágaó 'everyone'

The information in parentheses appears only in the phonological free form of the pronominal. (22d-f) only have free forms. Note that, with the exception of (22e) and (22f), these are all singular (cf. note 6). In fact, the only way to express plurality other than in these forms is periphrastically as in (23).

- (23) a. ti gixai pí -o kahá -p -í  
 1 2 also -obl go -imp -prox  
 'I and You go/We go.'
- b. gixai hi pío kahápí  
 2 3  
 'You and he go/You(pl) go.'

To keep the facts clear, this section is divided into three groups of phenomena (as in the introduction to this paper), labelled **reflexives**, **pseudotopicalization**, and **complex reference**, respectively.

### Reflexives

- (24) hi<sub>3</sub>(i) hi<sub>3</sub>(i/j) xib -áo -b -á  
 hit -telic -per -rem  
 (i) 'He hit himself.'  
 (ii) 'He hit him.'
- (25) gí<sub>2</sub>(i) gí<sub>2</sub>(i) xibáobá  
 'You hit yourself.'
- (26) ti<sub>1</sub>(i) ti<sub>1</sub>(i) xibáobá  
 'I hit myself.'

As was mentioned earlier, the verb form in these examples remains constant, since Pirahã verbs are not inflected for person, number, transitivity, reflexivity, etc. These examples might tempt us to conclude hastily that Pirahã merely lacks a morphological distinction between reflexives and nonreflexives. We suggest below, however, that the correct conclusion involves a different perspective on these pronominals and that reflexivity is best understood at a more abstract level of representation and thus is not relevant to the interpretation of these pronominals.

### Pseudotopicalization

The examples which follow are superficially similar to, yet fundamentally distinct from, topicalization in Pirahã, (cf. Sect. 4 below). For this reason, they are labelled as pseudotopicalization.



In these (and subsequent) sentences, the value of the subscripted indices are as follows:  $j \neq i$ ;  $k$  is free (i.e. it may take any antecedent in or outside of the sentence in question).

- (27) kohoibíhái<sub>(i)</sub> hi<sub>(i)</sub> ti<sub>(j)</sub> xibáobá  
 kohibíhái<sub>(i)</sub><sub>3</sub> hi<sub>(i)</sub><sub>1</sub> hit<sub>(j)</sub>  
 'Kohoibíhái hit me.'
- (28) kohoibíhái<sub>(i)</sub> hi<sub>(i)</sub> gí<sub>(j)</sub> xibáobá  
 'Kohoibíhái hit you.'
- (29) kohoibíhái<sub>(i)</sub> hi<sub>(i)</sub> hi<sub>(k)</sub> xibáobá  
 (i) 'Kohoibíhái hit himself.'  
 (ii) 'Kohoibíhái hit him.'
- (30) kohoibíhái<sub>(i)</sub> hi<sub>(i)</sub> xabagi<sub>(j)</sub> xibáobá  
 'Kohoibíhái hit xabagi'
- (31) \* kohoibíhái<sub>(i)</sub> ti<sub>(j)</sub> hi<sub>(k)</sub> xibáobá  
 \* 'Kohoibíhái I hit him/someone.'
- (32) \* kohoibíhái<sub>(i)</sub> gí<sub>(j)</sub> hi<sub>(k)</sub> xibáobá  
 \* 'Kohoibíhái you hit him/someone.'
- (33) \* Kohoibíhái<sub>(i)</sub> hi<sub>(j)</sub> hi<sub>(k)</sub> xibáobá  
 \* 'Kohoibíhái he/someone hit him/someone.'
- (34) \* Kohoibíhái<sub>(i)</sub> hi<sub>(j)</sub> xabagi<sub>(j)</sub> xibáobá  
 \* 'Xabagi hit Kohoibíhái.'

In (27)-(34), the relevant observations are that (a) the leftmost pronominal is obligatorily coreferential with the subject (as is especially clear in the starred examples; cf. also the forced, yet impossible, translation in (34)); (b) the rightmost pronominal is always free in reference (but see the next group of examples). Both occurrences of pronominals are optional, as seen in (35) (cf. also (17) and (18) above).

- (35) a. kohoibíhái xabagi xibáobá  
Kohoibíhái hit Xabagi.'
- b. kohoibíhái xibáobá  
(i) 'Kohoibíhái hit (someone).'
- (ii) '(Someone) hit kohoibíhái.'
- c. xibáobá  
'(Someone) hit (someone).'

As was mentioned, examples such as (27)-(34) are superficially reminiscent of a topicalization paradigm, as in (36):

- (36) a. John<sub>(i)</sub> hit { Bill,  
him<sub>(j)</sub>  
himself<sub>(i)</sub> }
- b. John<sub>(i)</sub>, he<sub>(i)</sub> hit { Bill,  
him<sub>(i)</sub>  
himself<sub>(i)</sub> }
- \*c. John<sub>(i)</sub>, I hit { Bill,  
him<sub>(j)</sub>  
himself<sub>(i)</sub> }

However, as will be shown in Sect. 4, this is not the case at all. Let us turn now to the final set of examples, what we refer to here as complex reference.

#### Complex reference

- (37) a. kohoibíhái<sub>(i)</sub> hi<sub>(i)</sub> xabagi<sub>(j)</sub> hi<sub>(j)</sub>  
xibáobá  
'Kohoibíhái hit Xabagi.'
- b.\* kohoibíhái<sub>(i)</sub> hi<sub>(j)</sub> xabagi<sub>(j)</sub> hi<sub>(i)</sub>  
xibáobá  
\* 'Xabagi hit Kohoibíhái.'
- c.\* kohoibíhái<sub>(i)</sub> hi<sub>(k)</sub> xabagi<sub>(j)</sub> hi<sub>(k)</sub>  
xibáobá  
\* 'Kohoibíhái someone hit Xabagi someone'

The ungrammaticality of (37b) and (37c) is due to the fact that, as in pseudotopicalization, the rightmost pronominal must have as its antecedent the rightmost NP while the leftmost pronominal takes the leftmost NP as its antecedent. A condition violated by the indices in (37b) and in (37c) when  $k \neq i/j$ .

- (38) a. kohoibíhái<sub>(i)</sub> xabagi<sub>(j)</sub> hi<sub>(i)</sub> hi<sub>(j)</sub>  
 xibáobá  
 'Kohoibíhái hit Xabagi.'
- b.\* kohoibíhái<sub>(i)</sub> xabagi<sub>(j)</sub> hi<sub>(j)</sub> hi<sub>(i)</sub>  
 xibáobá  
 \* 'Xabagi hit Kohoibíhái.'
- c.\* kohoibíhái<sub>(i)</sub> xabagi<sub>(j)</sub> hi<sub>(k)</sub> hi<sub>(k)</sub>  
 xibáobá  
 \* 'Kohoibíhái someone hit Xabagi someone.'

The explanation of ungrammaticality in (38b) and (38c) is the same as in (37b) and (37c).

In spite of its apparently greater complexity, (37) represents the most common type of sentence construction, clitic doubling of subject and direct object positions. As opposed to example (29), where the rightmost *hi* '3' is free, the rightmost *hi* '3' in (37) and (38) is obligatorily bound to the rightmost NP and is obligatorily disjoint from all other NPs in the sentence.

Any account of reference in Pirahã must offer a coherent treatment of these three sets of facts (and others which we will introduce below). Specifically, why are the pronominals free when no doubled NP appears yet obligatorily bound/disjoint in the manner shown above when a doubled NP is present? How are the correct referential "links" guaranteed? What is the structure of these examples?

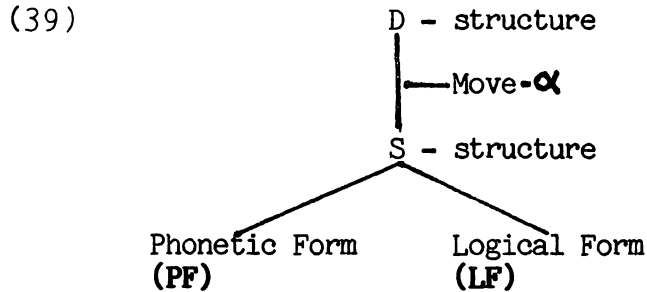
Before proposing what seems to us the most adequate analysis of these facts, let us briefly consider some basic principles of the model we will be working with.

### 3 The theoretical framework

While GBT is clearly a logical outgrowth of by now familiar research initiated in the late forties and early fifties into the nature of human grammatical capacity, UG, there are some rather fundamental differences between this model and previous versions, such as the well known **Aspects** framework (cf. Newmeyer (1980); Everett (1981)). We shall therefore discuss briefly these differences, concentrating on the areas of immediate relevance to the present study.

As in previous stages of generative grammar, GBT maintains that various grammatical levels and principles interrelate to generate a given sentence. We may diagram the theory as in (39) (cf. Fiengo

(1980), Chomsky (1981); van Riemsdijk and Williams (1981); and others for more details):



**D-structure** is the level at which  $\theta$  relations (cf. below) such as **agent**, **patient**, etc. are assigned. Its contribution to semantic interpretation beyond this is, in GBT, minimal - a sharp contrast with proposals in Chomsky (1965; 1971).

**Move- $\alpha$**  represents the transformational component in its entirety, literally allowing any syntactic category to be moved anywhere. Ungrammatical results are ruled out by general principles (cf. below) rather than ad hoc or overly specific structural descriptions for individual rules. **Move- $\alpha$**  may also apply at PF or LF (cf. Chomsky (1981); May (1977)). Further, categories moved by this rule leave a coindexed **trace** in the position from which they were moved.

**S-structure** is the interface between PF, LF, and the syntax. It is considerably more abstract than the **surface structures** of previous versions of generative theory since it contains traces and indices left by **Move- $\alpha$**  as well **empty categories**.

**PF**, the phonological component, will not concern us here. See Chomsky and Halle (1968); v.d. Hulst and Smith (1982), and others for some proposals.

**LF** is the linguistic input to interpretation. A clear outline of its basic features may be found in May (1977) and Chomsky (1981). We return to LF in the final section, as we test our analysis' predictions with regard to the effect of clitics on movement and interpretation in Pirahã.

The relations between the components of (39) are of two different types. First, the components are related in terms of input and output of rules - the **rule perspective**. Second, each sentence generated in (39) must comply with general conditions on well-formedness - the **systems/principles perspective**.

Although the nature of rules (morphological, syntactic, logical or phonological) continues to be an important topic of research in GBT, this second perspective which "... focuses on principles that

hold of rules and representations..." (Chomsky (1982:4ff)), might be fairly said to be the one that is most interesting as a source of insights into UG in current investigations. Those principles which will most concern us here are:

(40) **Government Theory:**

This is the pivotal system or subtheory in GBT, from which properties of most other systems are (at least in part) derivable. The basic intuition is that a lexical category (noun, verb, preposition, etc.) governs its complements. While there are many formal definitions of government in the literature, we will adopt the proposal of Chomsky (1982:19), although nothing crucial depends on this:<sup>7</sup>

- (41) " $\alpha$  governs  $\beta$  if  $\alpha = X^0$  (in the sense of X-bar theory),  $\alpha$  c-commands  $\beta$ , and  $\beta$  is not protected by a maximal projection."

A central notion in government theory is the **ECP (empty category principle)** which requires all empty categories except **PRO** (cf. below) to be governed.<sup>8</sup> next section.

- (42) **Case theory:** The principal intuition here is that there exist configurational and/or lexical requirements between nouns and the heads of phrases in which they occur which obligate these nouns or their traces to receive a syntactic Case (e.g. nominative, objective, etc.). These requirements are responsible for the **Case filter** which states that "..." every NP with phonological content must receive Case." (Chomsky (1982:6)).

For example, consider (43):

- (43) \* Who<sub>(i)</sub> was hit Bill by t<sub>(i)</sub>?

Among other problems in (43) is the fact that with **who** (or its trace) in subject position **Bill** is forced to remain in object position. But it is well known that passive verb forms do not assign Case to their objects. Normally, this would be remedied by raising **Bill** to subject position where it would receive nominative Case. Since the presence of **who** in (43) prohibits this, the Case filter is violated and the sentence is ungrammatical

- (44) **Theta ( $\theta$ ) theory:** The idea here is that predicates assign **thematic ( $\theta$ ) roles** to their arguments. The heart of the theory is:

the  $\theta$ -criterion:

- (45) (i) Every  $\theta$  role must be assigned to (just) one argument.  
 (ii) Every argument must be assigned (just) one  $\theta$  role.

To see how this works, consider (46):

- (46) a. John<sub>(i)</sub> was hit t<sub>(i)</sub> by Bill.  
 b.\* John<sub>(i)</sub> was hit Bill<sub>(j)</sub> by t<sub>(j)</sub>.

Since both **John** and **Bill** receive Case in (46), (**John** as subject, **Bill** through its trace), the problem of (46b) cannot be ascribed to Case theory. Now note that **Bill** receives the  $\theta$  role 'patient' from **hit** (and, arguably, the 'agent' role from **by**), while **John** receives no  $\theta$  role, since passives assign no  $\theta$  role to their subjects. Thus the  $\theta$ -criterion is violated twice - by **John** as an argument with no  $\theta$  role and **Bill** as an argument with two  $\theta$  roles.

- (47) **Binding theory:** Binding involves the relations between referentially dependent terms to their antecedents.

These dependent terms are of two kinds: (i) **anaphors**, those elements which have no inherent reference, requiring an antecedent in the immediate linguistic context (their **governing category**, cf. note 9 below); (ii) **pronouns**, dependent terms which may **not** have an antecedent in the immediate linguistic context, although they may take antecedents outside of this context. This subtheory reduces to two basic principles:

- A. An anaphor is bound in its governing category.<sup>9</sup>  
 B. A pronominal is free in its governing category.

A fuller typology of nominal categories (lexical and empty) is found below. These four sets of principles are central notions of GBT and crucial to the present study (cf. Chomsky (1981; 1982). Let us turn now to consider the concept and typology of empty categories in GBT.

The basic motivations behind the investigation of empty categories (**ecs**) are that: first, there exists clear evidence that **gaps** occur in the syntax where a syntactic category might have been expected to occur, as in (48)-(51):

- (48) Who did you see \_\_\_\_\_ ?  
 (49) John was hurt \_\_\_\_\_.  
 (50) Mary convinced her friend \_\_\_\_\_ to go to church.

- (51) Mas, quando ~~—~~ chegou, todos ficaram aliviados.  
 but, when arrived, all became relieved  
 'But when ((s)he) arrived, everyone was relieved.'

The second motivating factor is that these *ecs* have various properties with fundamental implications for the grammar as a whole.

Chomsky (1982:78) has proposed the following typology of the *ecs*; <sup>10</sup>

- (52) a. [ +anaphor, -pronominal ] 'trace'  
 b. [ +anaphor, +pronominal ] 'PRO'  
 c. [ -anaphor, +pronominal ] 'pro'  
 d. [ -anaphor, -pronominal ] 'variable'

(52b) and (52c) merit a bit more discussion. Recall from the Binding Theory that anaphors must be bound in their governing category while pronominals must be free in theirs. This leads to an apparent paradox since PRO must be both bound and free in its governing category. The solution is that PRO cannot have a governing category. One such environment is the subject position of infinitives, as in (50) above. (52c) is not available in English and other languages without the *pro*-drop parameter. It is in fact like any other pronoun except that it lacks phonological features. Languages with rich inflectional systems, such as Portuguese (cf. (51)) can easily "recover" the information lost in this lack of phonological content and therefore use *pro* frequently (subject to other restrictions which do not concern us here.)<sup>11</sup> With these basic notions in mind, then, let us turn to consider in more detail the analysis of the facts from Pirahã.

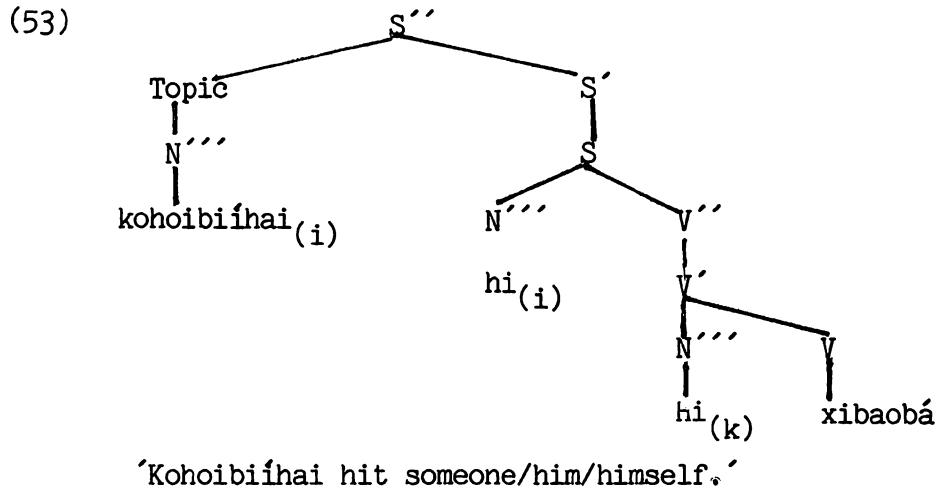
#### 4 Towards an analysis of intrasentential reference in Pirahã

##### 4.1 A failed analysis

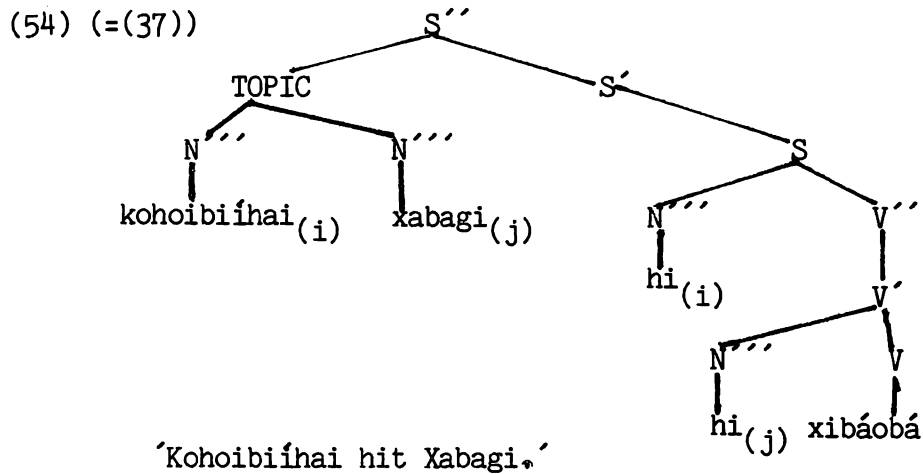
As was noted earlier, the facts grouped as pseudotopicalization and complex reference might appear to be merely a type of topicalization, in which the pronominals *ti*, *gí*, and *hi* are simply pronouns. Were this the case, it would be a gross error to analyze the facts as clitic doubling. Therefore, it is necessary to demonstrate conclusively that the facts here cannot be characterized as topicalization, before we can take up the issue of clitic doubling.

We showed previously that right/leftmost pronominals in the pseudotopicalization/complex reference data are obligatorily coreferential with the right/leftmost NPs, respectively. Under the topicalization hypothesis, these reference facts could perhaps be attributed to some sort of relationship between pragmatic prominence, linear restraints on language processing, and an analysis of topics in the spirit of, say, Reinhart (1982). But it is relatively easy to refute this hypothesis. As a first argument, consider the structures which might be required, assuming again that the pronominals are full

pronouns in NP position. A possible phrase marker for (29) would be (53):

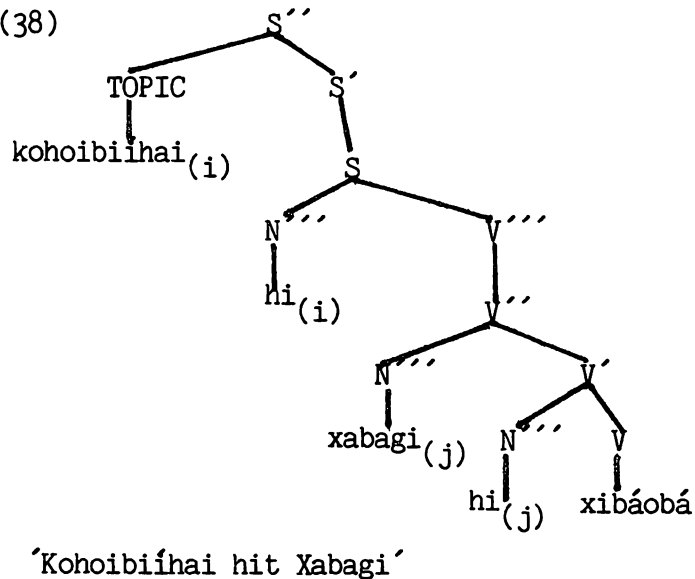


But now consider what sort of representation would be left us under this same hypothesis for (37) and (38).<sup>12</sup>





(55) (=38)



(55) is only one of many conceivable analyses under the TOPIC-pronoun hypothesis, but it is representative of the type of difficulty faced. Supposing (38) to be generated as (55), we create the serious problem of offering distinct treatments for each of the pairs, **kohoibíhai**<sub>(i)</sub>, **hi**<sub>(i)</sub> and **xabagi**<sub>(j)</sub>, **hi**<sub>(j)</sub>. In the first pair we apparently could derive the indicated coreference through some special rule of TOPIC interpretation along the lines of Chomsky (1977). But in the second instance, we would be forced to propose a different means of guaranteeing coreference. Further, we would need to decide exactly what sort of animal the N''' dominated by V'' is: argument (A) or nonargument (A') position - either answer being problematic for theory internal and language specific reasons. For example, if it is an A-position we have three such positions under VP and two with the same grammatical function (indirect object, under V''', as well as the V'' and V' direct objects). If it is an A'-position we have the unusual case of a nonargument position base generated under VP.

Alternatively, we could propose that **xabagi** is moved to its position in (55) in the PF component. This would resolve the configurational dilemma but would raise the more serious problem of how to restrict such a potentially powerful mechanism as PF movement of NPs, a problem which does not arise under the clitic analysis proposed below. In any case, it is clear that we are faced with some knotty configurational problems if we assume simultaneously that (37) and (38) are topicalized and that the pronominals are full pro nouns (cf. also the discussion examples (68)-(70) below).

Moreover, even if we could overcome the configurational problems, we would still be left with the problem of how to guarantee the correct reference relations. Surely, we cannot be satisfied with some mere muttering about a "pragmatic problem." We have no explanation for why the reference is assigned in the required fashion or how to

eliminate the **ungrammatical** (and not merely pragmatically anomalous) results of incorrect indexings. Finally, were all of these problems satisfactorily resolved, a much greater difficulty exists for a topicalization analysis of such examples - the fact that Pirahã has clearly topicalized structures (and in which structures such as (29), (37) and (38) can be embedded).

Note that, as they stand, pseudotopicalization and complex reference sentences lack any sort of special phonological or morphological marking which might distinguish them from other sentences. In general, we might reasonably expect some sort of device to be used (cf. Givón (1976)). And in Pirahã, there exists a series of examples in which such special marking **does** exist, what we consider to be real **topicalization**. For examples, consider (56)-(58):<sup>13</sup>

- (56)  $\begin{array}{ccccccc} \text{paigí}(i) & \text{hi}(i) & \text{xob} & \text{-áaxái} & / & \text{paigí} & \\ \text{Paigí} & \text{3} & \text{see} & \text{well} & & \text{paigí} & \\ \hline & & & & & & \end{array}$   
 '(As for) Paigí, he really knows a lot.'

- (57)  $\begin{array}{ccccccc} \text{hoaípi}(i) & \text{hi}(i) & \text{hoagí}(j) & \text{hi} & \text{xog} & \text{-i} & \text{-baí} \\ \text{Hoaípi} & \text{3} & \text{son} & \text{3} & \text{like} & \text{-ep} & \text{-int} \\ \hline & & & & & & \end{array}$   
 $\begin{array}{c} / \text{hoagí}(j) \\ \text{son} \\ \hline \end{array}$   
 '(As for) his son, Hoaípi really loves him.'

- (58)  $\begin{array}{ccccccc} \text{hi}(i) & \text{hi}(j) & \text{koho} & \text{-ai} & \text{-baí} & / & \text{páxaihi}(i) \\ \text{3} & \text{3} & \text{eat} & \text{-atelic} & \text{-int} & & \text{chicken} \\ \hline & & & & & & \end{array}$   
 $\begin{array}{c} \text{xigagí}(j) \\ \text{pepper} \\ \hline \end{array}$   
 '(As for) chickens and peppers, they really eat them.'

Note the separate intonational contours and pause between the main clause and its topic in these sentences. These structures are clearly highlighted as different from "run-of-the-mill" examples like (29), (37), and (38). Further, (58) shows clearly the difference in interpreting topicalized structures as opposed to pseudotopics or complex reference. For example, the only problem we create by reversing the indices in (58) is the rather difficult to imagine situation in which peppers eat chickens. Moreover, the subject **hi** in (58) could be substituted by **ti** '1' or **gi** '2' without affecting grammaticality:<sup>14</sup>

- (59) a.  $\underline{ti_{(i)} hi_{(j,k)} kohaibaí} / \underline{páxaihi_{(j)}}$   
 $\underline{xigagí_{(k)} (píai)}$   
 (also)  
 '(As for) chickens and peppers, I really eat them.'
- b.  $\underline{gí_{(i)} hi_{(j,k)} kohaibaí} / \underline{páxaihi_{(j)}}$   
 $\underline{xigagí_{(k)} (píai)}$   
 (also)  
 '(As for) chickens and peppers, you really eat them.'

Topic may also occur leftward in Pirahã, as (60) and (61) demonstrate:

- (60)  $\underline{xoogíái_{(i)}} / \underline{1\ ti_{(j)}\ 3\ hi_{(i)}\ xog\ -i\ -hiaba}$   
 $\underline{xoogíái_{(i)}} / \underline{1\ 3\ like\ -ep\ -neg}$   
 '(As for) Xoogíái, I don't like him.'
- (61)  $\underline{xahoaógíí_{(i)}} / \underline{xahoaógíí_{(i)}\ hi_{(i)}}$   
 $\underline{xahoaógíí_{(i)}} / \underline{xahoaógíí_{(i)}\ 3}$   
 $\underline{xisaxoi-baí}$   
 $\underline{sing-int}$   
 '(As for) Xahoaogíí, he sings a lot.'

Such examples are less frequent, however, than rightward topics.

Once again, while some reference possibilities in topicalized structures may produce a sensation of strangeness, or no change in acceptability whatsoever, such indexing changes in pseudotopics or complex reference produce clear ungrammaticality.<sup>15</sup>

To summarize, real topicalization differs from pseudotopicalization and complex reference in three ways: (i) real topicalization is phonologically marked, (ii) real topicalization allows coreferring R-expressions to occur in apparent violation of the BCs, and (iii) real topicalization is subject to a freer, more pragmatically oriented interpretation of reference in which acceptability judgements are much less sharp.

A topicalization analysis of (29), (37), (38) etc. would, therefore, fail to explain these contrasts. Also, it would have difficulty in deriving the configurations involved, and it is less satisfying empirically and theoretically than the alternative to be presented below.

Let us turn, therefore, to consider this alternative, beginning with a review of the notion of clitics in GBT.

## 4.2 An analysis of clitic doubling in Pirahã

**4.2.1 A note on the nature of clitics.** Kayne (1975:67) contrasts pronouns in French which can occur in environments where full NPs are allowed with those which may not. In so doing, he makes one of the first references to the term **clitic** in generative literature:

"Let us call the form of the pronoun that occurs in these environments its 'strong' form. In this class will fall **eux, nous, moi, toi, lui, elle, vous, elles**. Conversely, let us call the form of the pronoun that occurs... preposed to the verb its 'weak' or 'clitic' form, or simply 'clitic'. The direct object clitics corresponding to the above strong forms are **les, nous, me, te, le, la, vous, les**."

Borer (1981) is one of the several recent treatments of clitics which go beyond Kayne's analysis in arguing that clitics are fundamentally distinct from NPs and "strong" pronouns both in their generation and in their function (cf. (64) below). Consider, for example, her discussion of an important observation of Kayne's (p.50):

"R. Kayne has observed that...clitic-doubling constructions can only occur if the NP which is doubled is preceded by a preposition. This generalization (which Jaeggli calls "Kayne's Generalization") is accounted for by Chomsky (the Pisa Lectures), Aoun (1979), and Jaeggli (1980) by assuming that in clitic-doubling constructions the clitic... absorbs the Case features of the head..."

Borer thus argues in favor of the conclusion that (p.49): "...clitics are best characterized as part of the head constituent."

This conclusion will play a crucial part in the analysis of clitic-doubling in Pirahã which follows.

**4.2.2 Clitics, ecs and M-chains in Pirahã.** Before proposing the analysis which we believe best handles the data in question, we want to introduce some additional examples of clitic doubling in Pirahã. As these phrases demonstrate, clitic doubling is also possible in NPs and PPs in Pirahã.

- (62) a.  $xaxái_{(i)}$   $hi_{(i)}$   $kaií$   
 $xaxái_{(i)}$   $3_{(i)}$   $house$   
 'Xaxái's house'
- b.  $xigihí_{(i)}$   $hi_{(i)}$   $xíbaisi$   
 man  $3_{(i)}$  wife  
 '(The) man's wife'
- c.  $xigihí_{(i)}$   $hi_{(i)}$   $xíbaisi_{(j)}$   $hi_{(j)}$   $kaob -á$   
 man  $3_{(i)}$  wife  $3_{(j)}$  fall -rem  
 '(The) man's wife fell down.'<sup>16</sup>

Recall that nouns are Case assigners in Pirahã and that clitics are always optional (although most frequently present). Then the noun **kaií** 'house' assigns Case to the proper noun **xaxái** in (62a) with no need of a special preposition or suffix. The clitic's role in these examples will be made clearer below.

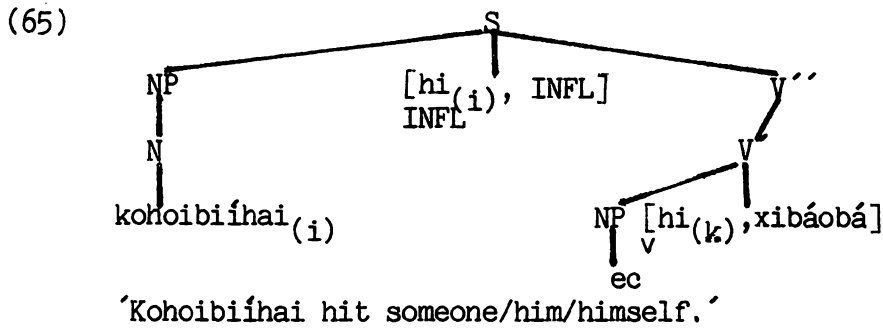
- (63) a.  $tábo_{(i)}$   $hi_{(i)}$   $xapó$   
 board  $3_{(i)}$  on  
 'on the board'
- b.  $baíxi_{(i)}$   $hi_{(i)}$   $giopai_{(j)}$   $hi_{(j)}$   $xigí -ó -xopí$   
 parent  $3_{(i)}$  dog  $3_{(j)}$  with -ob -go  
 'Mom/Dad went with the dog.'

To account for these new examples, as well as the reflexives, pseudotopics, and complex reference discussed earlier, let us supplement the categorial rules in (21) with the optional spell-out rule in (64) (cf. Borer (1981:52)).

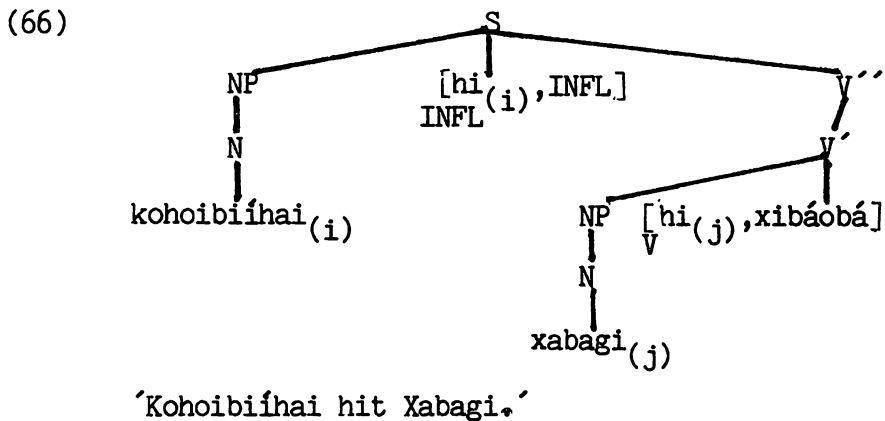
- (64)  $X_{[Case]} X \text{ --- } X_{[[Case, person], X]}$   
 $X = INFL, V, N, P$

Then, by (64), any category with Case features (e.g. INFL, nouns, verbs, and prepositions) may optionally realize these features as a clitic.<sup>17</sup>

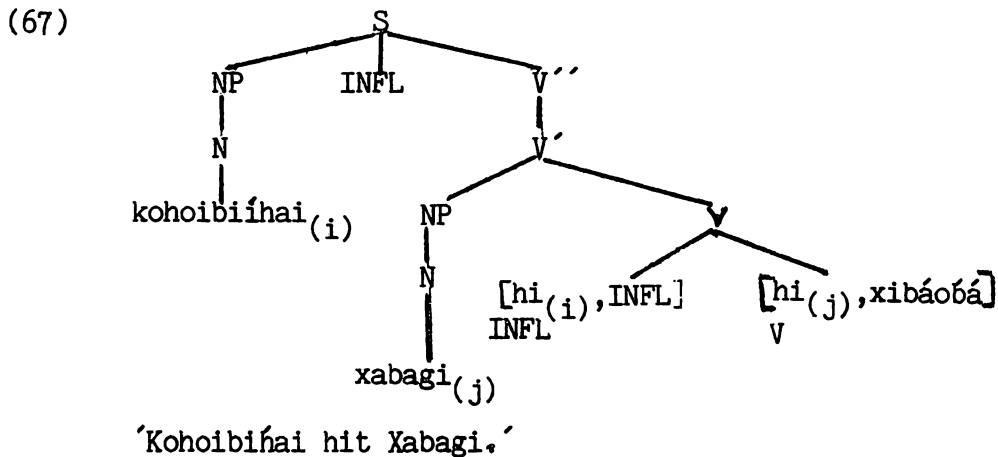
Consider, for example, the following structural analysis of the pseudotopicalized sentence of (29):



Now we may represent example (37) above as (66):



Then it is a simple matter to derive (38) from (37) through Move- $\alpha$ , producing (67):



We assume that this **INFL lowering** is in the PF component.

Note that we are not claiming here that **all** pronominals in Pirahã are clitics, only that the phonologically weak forms are. The present analysis allows pronouns to be base-generated normally under

NP. However, this raises a slight problem since then we have no clear explanation for the ungrammaticality of examples like (68) and (69):

(68) \* ti ti ti ti xibáob'a  
       1 1 1 1 hit  
       'I hit myself.'

(69)\* gíxai gí gíxai gí xibáobá  
       2 2 2 2  
       'You hit yourself.'

If, for example, the free form of **ti** '1' (which appears in examples like (i) Speaker A: **kaoi kaipí** 'Who did it?' Speaker B: **ti** 'me') is generated under NP and the bound **ti** as a feature spell-out on INFL and/or V, then (68) should be grammatical, likewise for (69). Although I have no really satisfying explanation here, it is likely that (68) and (69) are ruled out since they would be incredibly redundant semantically and wierd phonologically.

On the other hand, our analysis explains examples such as (70) quite handily:

(70) ? hiapióxiái<sub>3</sub>(i) hi<sub>3</sub>(i) hiapióxiái<sub>3</sub>(j) hi<sub>3</sub>(j)  
       xibáobá.  
       hit  
       'Someone hit someone.'  
       He him himself

(70) seems to be less redundant sematically and phonologically due to the vagueness of **hiapióxiái** and **hi** as well as their large phonological difference, and therefore no unacceptability is created by the generation of this type of example. Moreover, insofar as a topicalization analysis would assume both **hiapióxiái** and **hi** to be full pronouns, (70) seems to be a rather conclusive counterexample.

It is further assumed here that binding by clitics produces no violations of the BCs since, presuming that only binding falls outside the scope of Binding theory (in other words, that clitic binding is of a different sort, such as Bok-Bennema's (1981) **M-binding**). We return to this rule of INFL-lowering directly. First, we need to discuss just how to guarantee the correct indexing of clitics and doubled NPs, including the **ec** in (65).

The reader familiar with Stowell (1981) and Borer (1981) will have noted that our formula  $y[cl, X]$  is a simplication. In the works just cited, it is argued tha X is in fact a bundle of information (cf. Borer (1981:54ff)). Among these bits of information

are the  $\theta$ -role assigned by the head, the clitic and its index, and a slot for the index of the **complements** of X (cf. the works cited for further details). Then we will have something along the lines (71) (cf. Borer (1981:55)).

$$(71) \quad \left[ \begin{array}{c} \theta_1 \\ \Delta \\ cl_i \end{array} \right]$$

Then, presuming in (68) to be the slot for the index of X's complements, we may make the reasonable assumption that the index of the clitic and in (71) must match or the structure will be ruled out. Under this more detailed account, correct coindexation between clitics and their doubled NPs is then derived from this **complement matching requirement**.<sup>18</sup>

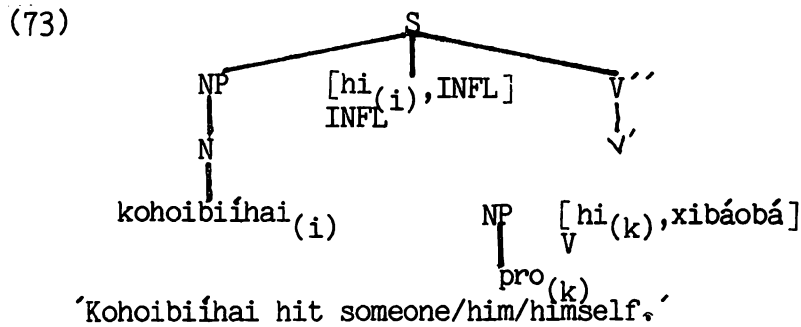
But at this point we come to a major difference between Pirahã clitics and those studied in these other works. The difference is that since no special Case assigner is necessary, it would appear that clitics in Pirahã do **not** absorb Case but in fact **transmit** Case to their doubled NPs, just as the bare categorial head on which they appear would, were they not present. I submit that this transmission is done through coindexation, reminiscent of transmission of Case and/or  $\theta$  roles through traces in syntactic chains (cf. Chomsky (1981:333) and Safir (1982)). But since such chains either involve two elements in argument positions (as in NP movement) or are headed from a nonargument position, with a trace in argument position (as in WH-movement), it seems that what we are faced with here is a different type of chain. Let us call it for the moment a **M (morphological) chain**. We return to discuss this type of chain and the parameters and predictions it involves below.

The most urgent task facing us at present is the identification of the **ec** in (65). We will adopt recent proposals (Chomsky (1981; 1982); Safir (1981) that identify **ecs extrinsically** through the following criteria (where **F** is some set of grammatical features. Cf. Chomsky (1981:330)):<sup>19</sup>

- (72) (i)  $\alpha$  is a **variable** if it is locally A'-bound and in a A-position.  
 (ii)  $\alpha$  is **pronominal** if  $\alpha =_{NP} [F, (P)]$  where P is a phonological matrix and F is non-null, and either (a) or (b):  
 (a)  $\alpha$  is free  
 (b)  $\alpha$  is locally A bound by  $\beta$  with an independent  $\theta$  role.  
 (iii) if  $\alpha$  is an empty category which does not fall under (i) or (ii) and is locally A-bound, then it is an **anaphor**.

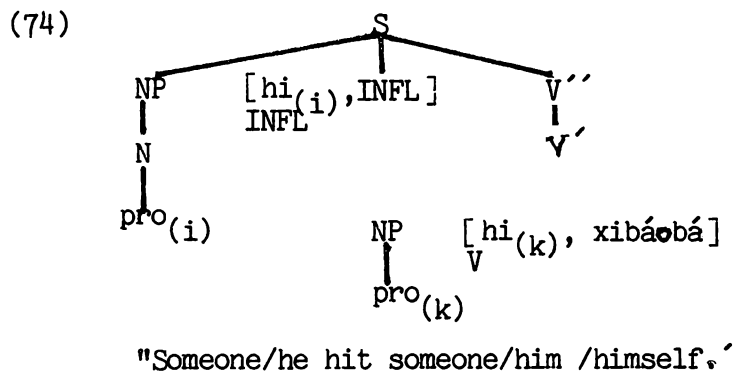


Now, following Borer (1981), we suppose that the **ec** in (65) is governed by the V clitic, **hi**<sub>(k)</sub>. Thus, it cannot be PRO (since, as was mentioned earlier, PRO cannot be governed). But the **ec** could not be [+anaphor, -pronominal] either since, arguably, this would entail the formation of a syntactic chain with its antecedent, violating the  $\theta$  criterion (cf. Chomsky (1981:333) and Safir (1982)). Further, the **ec** cannot be [+anaphor, -pronominal], a variable, since it is not locally A'-bound, as required by (72). Since this **ec** is free and neither a variable, an anaphor or PRO it must be **pro**, a pronoun without phonological realization. Under this assumption, the structure of (29) would be more accurately represented as (73):



Then, in (73), the index **k** on **pro** and the **hi** in V may be interpreted in LF as either equal to **i** on **kohoibifhai** and **hi** in INFL, producing a reflexive translation, or as distinct from **i** producing a nonreflexive translation. But since the BCs require pronominals to be free in their governing category, the governing category for **pro** in (73) cannot be S or it would violate the BCs under the reflexive interpretation. Recalling our earlier definition of governing category, let us assume that the clitic in V may serve not only as a governor for the **ec** but as a **SUBJECT** accessible to it. Then the governing category for **pro** in (73) will be V''.<sup>20</sup> The governing category for the subject of S will be S by virtue of the clitic/features in INFL.

Now let us consider the structure of reflexives, as examples (24) above, represented as (74).



Thus, under the present analysis, the fact that there are no special morphological forms for reflexive pronominals is explained, as is the ambiguity of such constructions, straightforwardly by the fact that reflexivization is a case of clitic doubling, involving the pair [pro, cl]. But now we must ask more specifically what kind of relationship obtains between the clitic and its doubled NP.

Chomsky (1982:87ff) assumes that the clitic may form a chain with the doubled element but that the clitic cannot transmit Case in such chains (citing references already mentioned in this paper on the "absorption" of Case by the clitic). However, it is not obvious that any such chain exists for the languages Chomsky considers. Adopting Borer's complement matching proposal, the doubled NP receives its  $\theta$ -role directly from the  $\theta$ -role directly from the  $\theta$ -slot on the verb (cf. (71) above), coindexation with the clitic being an independent requirement. In fact, proposing that clitics and NPs/*ecs* form a **syntactic** chain, in the sense of Chomsky (1981:333) would produce a rather curious type of chain which the lower member, the clitic, can **not** transmit Case. Any proposed syntactic chain would thus be deficient in this respect. Our proposal is that the failure of Case transmission is explained by the fact that no chain exists in these languages,  $\theta$ -role transmission being a function of the complement matching requirement and " $\theta$ -slots" on the verb.

On the other hand, in languages like Pirahã, where doubled NPs do receive features of Case,  $\theta$ -roles, etc. through the clitic, it seems that a further concept is needed, what we referred to earlier as **M-chains**. In Pirahã, direct objects, subjects, genitives and postpositional objects may receive their required features from V, INFL, N, and P, respectively, by entering into an M-chain with their coindexed clitic.

Let us define an M-chain by the following.

- (75)  $MC = (\alpha, \beta)$  is an **M-chain** iff:  
 (i)  $\alpha$  is a morphological category on  $X^0$   
 (ii)  $\beta$  is an argument of  $X^0$   
 (iii)  $\alpha$  **M-binds**  $\beta$  <sup>21</sup>

Then, for any M-chain  $(\alpha, \beta)$ ,  $\alpha$  transmits the relevant features of  $X^0$  to  $\beta$  (cf. Borer (1981) for more discussion of the features involved). That is, in languages with M-chains (and not merely M-binding), Case,  $\theta$  roles, etc. are assigned to these and not simply absorbed by the morphological category,  $\alpha$ .

We might propose the utilization of M-chains to explain the properties of clitic doubling in Pirahã as opposed to cases such as Hebrew and River Plate Spanish which do not allow M-chains, thus requiring a separate Case assigning device, since the clitic absorbs the Case of the phrasal head but it has no means to transmit it.

Clearly, the hypothesis that clitics form M-chains with their doubled NPs makes different empirical predictions than theories which do not recognize this possibility. Borer (1981) gives some interesting evidence from Hebrew which supports her contention that clitics absorb but cannot transmit Case. We will state this argument here and then test its predictions with regard to Pirahã. Consider the contrast in Hebrew between direct questions (76a) and free relatives (76b).

- (76) a. \*  $ma_i$  xasavti 'al -av<sub>i</sub> ec<sub>i</sub> (Borer's (138b))  
 'What did I think about?'<sub>i</sub>
- b.  $ma_i$  se-xasavti 'al -av<sub>i</sub> ec<sub>i</sub> (Borer's (138a))  
 'Whatever I thought about,'<sub>i</sub>

Borer explains this contrast by assuming (i) the extrinsic definition of **ecs** in (72) above, by which the **ec** in (76) is a variable, and (ii) that an **ec** is a variable if it has Case. Then she further argues that in free relatives, but not in interrogatives, the fronted WH element receives Case through Case marking into COMP, and therefore the **ec** may be said to have Case by virtue of being in a syntactic chain with the fronted element. Thus, (76b), a free relative, is grammatical.

But what about (76a)? Since the **ec** is A'-bound it should be a variable. Yet, because the clitic cannot transmit Case to this **ec** and since the fronted WH element cannot receive Case through COMP as its free relative counterpart in (76b) could, the requirements on variables produce contradictory specification of the **ec**, ruling (76a) ungrammatical.

However, according to our predictions here, if the clitic in (76) could form an M-chain with the doubled NP and/or its trace, the **ec** would fully satisfy the extrinsic definition of a variable, by receiving Case, and the structure would be grammatical. This would mean that in Pirahã, as opposed to Hebrew, WH movement should apply freely in clitic doubled constructions. Thus interrogatives and wide-scope interpretation in LF (cf. May (1977)) would be possible in clitic doubled configurations. Although at present we have no clear data on wide-scope vs. narrow-scope interpretation in Pirahã, there is abundant evidence from interrogatives.

Interrogatives in Pirahã are generally formed through verbal affixes. However, there does exist a free form WH element, **kaoi**, corresponding to English **who**. This form appears in constructions such as (77) and (78):

- (77) xaoó<sub>i</sub> hi<sub>(i)</sub> koi<sub>(j)</sub> hi<sub>(j)</sub> kob -ai híx  
 foreigner 3 who 3 see -atelic inter  
 'Who does the foreigner see?'

- (78)      kaoi<sub>(i)</sub> hi<sub>(i)</sub> ti<sub>(j)</sub> xaho -ai      -xiig -á  
             who<sub>(i)</sub> 3<sub>(i)</sub> 1<sub>(j)</sub> speak -atelic -cont -rem  
             'Who is speaking to me?'

Adopting the standard GBT analysis of interrogatives, *kaoi* 'who' will move into COMP at LF.<sup>22</sup> Thus, we will derive LF structures:

- (79) [<sub>S'</sub> kaoi<sub>(j)</sub> [<sub>S</sub> xaoói<sub>(i)</sub> hi<sub>(i)</sub> t<sub>(j)</sub> kobai híx]]

- (80) [<sub>S'</sub> kaoi<sub>(i)</sub> [<sub>S</sub> t<sub>(i)</sub> hi<sub>(i)</sub> ti<sub>(j)</sub> xahoaxiigá]]

According to the present analysis, these examples are grammatical since the clitic forms an M-chain with the trace of WH movement thus transmitting Case (θ roles, etc.) to it. The alternative analyses which do not recognize the possibility of such Case transmission by clitics via M-chains wrongly predict (79) and (80) to be ill formed LF structures. It cannot be objected that no movement takes place here since *kaoi* is quite clearly a WH element in Pirahã and must be raised in LF due to the very nature of WH interpretation in GBT. And this is true whether or not the language in question has movement in the syntax (e.g. English) or not (e.g. Chinese, cf. Huang (1981)).

It therefore seems clear that the facts here require us to recognize a new parameter of clitic configuration in UG - that of M-chains wherein clitics not only receive or express Case and other features of the phrasal head but transmit them to their doubled NPs.

One final question which needs to be answered before concluding the present study involves the nature of INFL in Pirahã, specifically, why is INFL-lowering into V'' optional in Pirahã but required in (most) other languages?

Recall that Pirahã does not express tense. Therefore, if we assume that the approximately sixteen basic aspectual distinctions in Pirahã (cf. Everett (to appear)) are generated directly on the verb, then INFL in Pirahã is primarily nominal in nature.

According to Safir (1981:427), the rule INFL-lowering is to be explained in terms of a **tense filter**:

- (81)      "The Tense Filter:  
             Tense features must be spelled out on a verbal  
             phonological base."

If this filter is in fact the primary motivation for the obligatory application of INFL-lowering (Chomsky's (1981:256ff) "rule - R") then we might expect that in a language without tense in INFL,

such as Pirahã, the application of this rule is optional, which is in fact the case.

Thus we have explained the wide array of clitic doubling configurations in Pirahã and how these configurations differ from those in more well known languages.

At this point it seems reasonable to ask why languages should differ along these lines and, specifically, what makes clitics so different from other nominals. We leave this more speculative discussion for the appendix.

### **Appendix: Clitics, pronouns and affixes**

Typologists have long been intrigued with how agreement features on verbs and other categories develop. Specifically, various works have raised the question of how such features relate synchronically and diachronically to pronominals. In this appendix, I would like to offer a few suggestions as to how the facts observed by such typological research might be incorporated into a formal model such as GBT. Givón (1976) has proposed that pronouns are regularly "reanalyzed" diachronically as verb agreement, on the basis of examples such as (82):

- (82) a. The man, he came.  
       b. The man he-came.

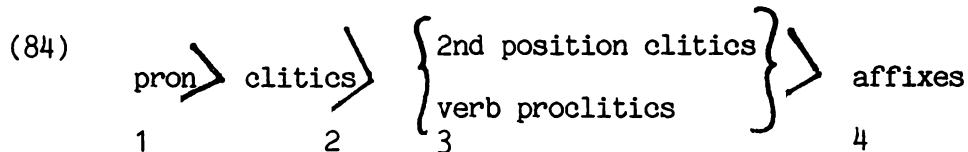
According to Givón, speakers eventually may come to use the marked, topicalized structure in (82a), as an unmarked, nontopicalized sentence, with no pause following the first constituent and with the pronoun phonologically bound to the verb. Over the course of time, so this reasoning goes, the phonologically weak pronoun may come to be reanalyzed as a verbal affix, losing its status as an independent word.

Similar suggestions are to be found in other works such as Shaul (1983) and the references cited there, in which "diachronic stages" are proposed, as in (83):



It is worth asking how this important work by typologists is to be understood in a theory such as GBT.

I submit that elements in (83) are defined by their relation to specific parameters of UG selected by the language in question. As a point of departure, let us propose a more inclusive version of (83), (84):



Exactly how do (1)-(4) in (84) differ? On the one extreme are the pronouns. Within GBT these elements can be further subdivided into pronominals and anaphors, as we have already seen. However, as a class, the factors which seem to best distinguish pronouns from other entities of (84) are their (i) phonological independence of the verb and (ii) inability to transmit syntactic features to their antecedents. Clitics in position 2 (**2 clitics**) then differ from pronouns in being more phonologically dependent on the verb and less restricted in their syntactic positioning, occurring pre- or postverbally and in doubled constructions in many languages (e.g. River Plate Spanish, Hebrew, etc. cf. also Shaul's discussion of Tepiman). However, **2 clitics**, like pronouns, still cannot transmit syntactic features to their doubled NPs.

But turning to the elements of position (3) in (84), **3 clitics**, we notice more rigid syntactic positioning than **2 clitics**. For example, according to Shaul (op cit:259), subject clitics in Tepiman went from relatively free order to more rigid positions:

"In the older Tepiman data, word order of nominals and predicate is rigidly SOV. The subject clitics, however, are movable (object pronominals being verb proclitics). In the modern languages, however, word order tends to be free while the positioning of the subject clitics is fixed."

A possible explanation for this **2 clitic** vs **3 clitic** distinction is the parameter of M-chains. Let us suppose that a language may choose to allow its clitics (all or some) to enter into M-chains with arguments or, in other terms, to allow Case,  $\theta$  roles, etc. to be transmitted via clitics. Then it must "reanalyze" these pronominals as a morphological, nonargument category (or the  $\theta$ -criterion is violated). Note that this is only a question of logical, not chronological order - we are dealing with parameters, not functional explanations). This could explain why such elements are more tightly ordered - as morphological categories they are rooted/fixed in specified morphological slots, e.g. in INFL and V, explaining second position and verb proclitics.<sup>23</sup>

The final "state" involves allowing both M-chains and the complete morphological absorption of these elements by the verb, etc. This last step is really quite natural, according to our view, since once M-chains are allowed, clitics merely fulfill the syntactic role of affixes. If they have no other role (pragmatic, semantic, etc.) then they lose their categorial distinctiveness and a natural step would be to simply absorb them into the verb morphology (something

which seems close to occurring in Pirahã and which has happened in other languages).

Note, too, that this speculation in fact makes some rather easily testable empirical predictions, supposing that the theory of clitics in Borer (1981) and the notion of M-chains suggested here are correct. **3 clitics** should, just as affixes, allow WH-extraction from doubled constructions as well as **quantifier raising (QR)** from such positions. This is due to the fact that the **ecs** left in such positions will receive Case and may be interpreted without contradiction as variables (cf. (79) and (80) above). Further, **3 clitics** should eliminate the need for prepositions or other Case assigners for the doubled NP. We have shown Pirahã clitics to be of this type.

In Everett (1983b), it is suggested that Piedmontese, a Romance dialect, presents further evidence supportive of this. For example, clitic doubling is obligatory (except from direct object position in non-topicalized structures. Cf. the work just cited for some discussion) and requires no special Case assigners (although a preposition is **optional** with dative Case):

(85) a. mi j u purta -je  
1 pronoun 1-clitic have brought -dative clitic

al liber al dzyzép  
the book the Joseph

'I bought the book to Joseph.'

b.\* mi u purta -je al liber al dzyzép

c.\* mi j u purta al liber al dzyzép

(86) a. {al} liber, j u purta -j  
{an}

(the) book 1-clitic have brought -dative clitic

-lo al dzyzép  
-accusative the Joseph  
-clitic

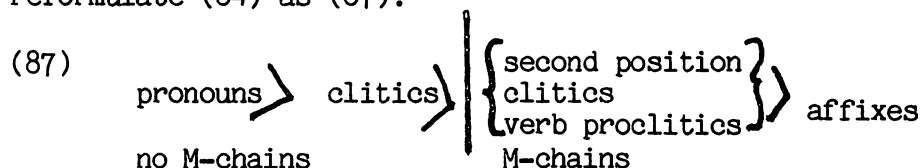
'(As for) the book, I have brought it to Joseph.'

b.\* {al } liber, j u purta -je al dzyzép  
{an }

(85) and (86) show clearly that clitic doubling is not only obligatory in Piedmontese but that no special Case assigner is necessary. Moreover, R. Ilari, a native speaker of Piedmontese, tells us that (86) allows either a definite or indefinite element in Topic

position. If we presume that, following May (1977), QR applies in LF to interpret indefinites then we must conclude that the clitic *lo* in (86) may form an M-chain with the *ec* argument of *purta* 'brought' to attribute Case to the variable left at LF by QR. Piedmontese lends some support to our hypothesis.

If in fact this reasoning is on the right track, we may reformulate (84) as (87):



Thus, we will have taken a first step towards explaining the observations of typologists suggested by (83) in terms of testable/falsifiable universal parameters of UG.

### Notes

\* Pirahã is a member of the Múra language family of central Brazil which also included the now (probably) extinct dialects of Múra, Bohurá, Yaháhi, Torá, and possibly, Matanawi (although data are scarce). The phonemes of Pirahã are /p/, /t/, /?/, /b/, /g/, /s/, /h/, /i/, /a/, /o/. /?/ is represented orthographically as 'x'. Pirahã also has two (register) tones, high and low represented orthographically (on vowels) as /' / and zero respectively.

Abbreviations used in this paper are (syntactic category labels capitalized):

- cl - 'clitic'
- COMP - 'complementizer node'
- compcert - 'complete certainty evaluative'
- cont - 'continuative aspect'
- DET - 'determiner'
- ep - 'epenthetic'
- imp - 'imperfective'
- INFL - 'inflection'
- int - 'intensive'
- INTER - 'interrogative'
- iter - 'iterative'
- MOD - 'modifier'
- neg - 'negative'
- obl - 'oblique'
- perf - 'perfective aspect'
- prox - 'proximate aspect'
- punc - 'punctiliar aspect'
- relcert - 'relative certainty evaluative'
- rem - 'remote aspect'
- 1 - 'first person singular'
- 2 - 'second person singular'
- 3 - 'third person singular'



The writing of this paper was supported in part by a fellowship from the Americal Council Societies partially funded by a grant from the US, Department of Health, Education, and Welfare.

1. Throughout this paper, subscripted letters indicate reference relationships.

2. These terms will be explained more fully in Sect. 3 below.

3. Other readings less relevant to the present context are:

$$\left. \begin{array}{l} \text{Someone} \\ \text{He} \end{array} \right\} \text{ hit } \left\{ \begin{array}{l} \text{someone} \\ \text{him} \\ \text{himself} \end{array} \right.$$

4. The (b) examples are grammatical with pause after the first constituent giving the respective readings:

(i) 'Coffee, Kohoihihai drinks a lot of.'

(ii) ?'Everything, God made.'

The reader is referred to Sect. 4.1 for further discussion.

5. This presumes that Pirahã is a **configurational** language. See Hale (1983) for a clear exposition of the differences between configurational and **non-configurational** languages.

6. This system may be said to violate Greenberg's (1966:96) **Language Universal 42**:

"All languages have pronominal categories involving at least three persons and two numbers."

It is not really clear whether (22e) and (22f) should be treated as pronouns per se or simply as particles. (22f) at times is translatable as "all of it" and it is clearly a compound form. (22e) may be translated as 'let us' (+ verb) or as 'C'mon'. Thus, taking (22a)-(22d) as the clear cases, this system is the simplest yet documented to my knowledge. Moreover, an alternative form of the third person pronoun, ?i, seems to be correlated in a high percentage of cases with feminine gender. While we have nothing more to say on this here, this would, if correct, violate still another language universal, Greenberg's (1966:95) **Universal 36** (since Pirahã would then have gender but no number distinctions): "If a language has the category of gender, it always has the category of number."

7. **C-command** may be defined by: c-commands iff neither dominates the other and the first branching node which dominates dominates. A **maximal projection** is the largest expansion of a lexical category X allowed by X' theory (and the language in question).

8. More exactly, these are required to be **properly governed**- This

will not concern us here, however, except to say that clitics (following Borer (1981)) properly govern their coindexed positions, permitting **ecs** in these positions.

9. Governing category may be defined as (following Huang (1983:557); cf. also Reuland (1983:127ff)):

" $\alpha$  is a governing category for  $\beta$  if and only if  $\alpha$  is the minimal category containing  $\beta$ , a governor of  $\beta$ , and a **SUBJECT** that, if  $\beta$  is an anaphor, is accessible to  $\beta$ ,"

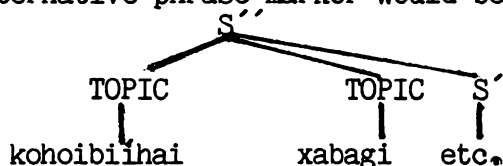
I will have more to say of **SUBJECT** in Pirahã below. See Chomsky (1981:209) for a definition of **SUBJECT**. Basically it is the most prominent nominal element in a particular category.

From other principles, we may also derive the fact that nouns with inherent reference, **R(eferent)-expressions**, (e.g. 'John', 'the dog', etc.) may not be bound.

10. These features represent the distinctions possible between **ecs** as well as **lexical** (i.e. phonologically realized) NPs. As is noted in Sect. 4, however, the **ecs** of a given structure are determined **extrinsically**, i.e. by their behavior in that particular environment, rather than **intrinsically**, i.e. inserted predefined as (52a)-(52c).

11. But such languages cannot be limited to just those with agreement inflected verbs, since Pirahã has prodrop without this type of verbal morphology. See below for an analysis of some prodrop constructions in Pirahã.

12. An alternative phrase marker would be:



This seems less desirable, however, if we accept Kayne's (1981) suggestions on restricting tree structures. Note, too, that this topicalization hypothesis could be revised by analyzing the pronominals as clitics, along the lines of our suggestions below, to eliminate some of these configurational difficulties. But, as we show, this still leaves insurmountable problems of other types.

13. The symbols used are:

/ = pause;    ˆ = rising intonation

14. The reader may be puzzled by the multiple indices associated with **hi**. In such cases, **hi** refers to both topics, simply lacking a

morphological distinction for plurality. In Everett (1983b), I argue that multiple indices are necessary in a system without **anaphoric indices** (in the sense of Chomsky (1980)), to account for **split antecedents**. These arguments are largely superseded, however, by Higgenbotham (1983).

15. In light of the Binding Theory discussed in Sect. 3, these examples need some explanation. It would appear that the proper nouns in examples like (61) are in violation of the BCs, which prohibit binding of such R - **expressions** in all environments. The best answer seems to be that in Pirahã (and I would expect other languages), S'' (but not S') is simply outside the Binding theory's domain, more a discourse type of interpretation strategy. (Cf. Everett (to appear) for discussion of Pirahã discourse). For example, (i) is hopelessly ungrammatical (with or without clitics. Cf. this with (61)):

- (i) \* $\left[ \begin{array}{c} \text{kóxoí} \\ \text{S'} \\ \text{kóxoí} \end{array} \text{ kóxoí} \text{ xibáobá} \right]$   
       kóxoí kóxoí hit  
       'Kóxoí hit Kóxoí.'

16. Note the structural ambiguity in (62c). This example may be parsed as either (i) or (ii), with the semantic differences noted:

- (i)  $\left[ \left[ \begin{array}{c} \text{xigihí} \\ \text{S'} \\ \text{NP} \end{array} \left[ \text{hi} \text{ xíbaisi} \right] \right] \left[ \begin{array}{c} \text{hi} \\ \text{VP} \end{array} \text{ kaobá} \right] \right]$   
       'The man's wife fell down.' (as in (62c))'
- (ii)  $\left[ \left[ \begin{array}{c} \text{xigihí} \\ \text{S} \\ \text{NP} \end{array} \right] \left[ \begin{array}{c} \text{hi} \\ \text{INFL} \end{array} \dots \right] \left[ \begin{array}{c} \text{xíbaisi} \\ \text{VP} \end{array} \text{ hi} \text{ kaobá} \right] \right]$   
       'The man threw his wife down.'

17. Note that the explanation of clitic configurations given below, while concentrating on phonologically realized clitics is equally valid when the clitic is not phonologically realized, that is, when (64) does not apply. When the clitic is not realized phonetically, its empty slot on the phrasal head may still form an M-chain with the doubled NP or **ec**.

18. It is clear that guaranteeing the correct indexation of clitics in INFL cannot be complement matching in this sense, since INFL has no complements (meaning arguments to which it assigns a  $\theta$  role). A first suggestion would be to understand the indexing of INFL with [NP,S] along the lines stipulated in Chomsky (1981:211): "AGR is coindexed with the NP it governs."

In this and subsequent work, Chomsky assumes the expansion of INFL to be:

- (i) INFL  $\longrightarrow$  [+AGR, + Tense]

AGR is a nominal feature complex. Since tense, as has been shown, is not a feature of INFL in Pirahã, and since we assume aspect to be generated on the verb (although this is not crucial were aspect in INFL we would simply modify the structures involved in the obvious way), then INFL as we are using it here for Pirahã is in fact AGR. The clitic in INFL (AGR) is the spell-out of the nominative Case (and agreement) features. This does not substantially alter the present analysis, however.

19<sup>i</sup>. Local X binding is defined as (Chomsky (1981:184ff))

- (i) "...  $\alpha$  is X - bound by  $\beta$  if and only if  $\alpha$  and  $\beta$  are coindexed,  $\beta$  c-command  $\alpha$ , and  $\beta$  is in an X - position..."
- (ii) "...  $\alpha$  is locally bound by  $\beta$  if and only if  $\alpha$  is X - bound by  $\beta$ , and if  $\gamma$  - binds  $\alpha$  then either  $\gamma$  Y - binds  $\beta$  or  $\gamma = \beta$ ."

In these definitions X/Y refer to binding from an argument (A) or nonargument (A<sup>0</sup>) position.

20. Considering V'' as a governing category in Pirahã may not be so unusual as it appears at first sight. The primary difficulty would be to establish that V'' has a **SUBJECT** accessible to the anaphor or pronominal *ec*, etc. But it seems to me that the clitic is an accessible **SUBJECT** for the [NP, V'] position, since it must agree with this NP. In fact, the common definitions given for **SUBJECT** hold quite well for V'' in Pirahã. Reuland (1983:127) proposes that **AGR** is "... indeed a **SUBJECT** in the sense required by the binding theory." Since the verb *does* have (object) agreement, independent of the (subject) agreement of INFL, then it seems reasonable to view this (clitic) agreement in V as a possible **SUBJECT**.

This would in effect make V'' a subtype of (or semi) clause. If this were the case, we might expect to find some other clause-like properties of V''. Such properties do seem to exist. Consider, for example, (co)relative clauses:

- (i)  $\left[ \begin{array}{c} \text{ti} \\ \text{S} \end{array} \text{ baósaápirisi} \text{ xog} \text{ -abagai} \left[ \begin{array}{c} \text{páxai} \\ \text{S} \end{array} \text{ hi} \right. \right.$   
 1 hammock want-frustrated priest 3  
 initiation  
 $\left. \left[ \begin{array}{c} \text{go} \\ \text{V}'' \end{array} \text{ -ó} \left\{ \begin{array}{c} \text{baósaápirisi} \\ \text{ec} \end{array} \right\} \text{ bag} \text{ -áo} \text{ -b} \text{ á} \right] \right]$   
 WH -obl sell -telis -perf -rem  
 'I want the hammock which the priest sold.'

Notice that these clauses may have an overt NP or an *ec* in the embedded relativized position. The interesting fact in this context, however, is the WH element *go* (used in interrogatives also, cf. Everett (to appear)). This *go -ó* appears to the right of the INFL clitic *hi* in the subordinate clause and is marked with the oblique suffix *-o*, two clear evidences that it is in  $V''$ , at the far left periphery. A possible explanation for *go* being in  $V''$  rather than *S* may be found in considering  $V''$  as a governing category (although, admittedly, this is quite speculative). The fact that the overt NP may co-occur with *go* shows too that *go* could not be simply a WH element *in situ*, but that it is a type of semicomplementizer at the  $V''$  periphery. We will not pursue this further here, noting merely that  $V''$  in Pirahã does have some peculiar properties which might be partially explained by its status as a governing category.

21. Bok-Bennema (op cit) defines M-binding as:

"M-binding:  $\alpha$  M-binds  $\beta$ , if  $\alpha$  is coindexed with  $\beta$ , and  $\alpha$  c-commands  $\beta$ , and  $\alpha$  is a morphosyntactic category and  $\beta$  is an argument."

22. To take a more neutral stance, since COMP seems to be rightward (if relevant at all) in Pirahã, while WH movement is leftward in S((79) and (80) above) or in  $V''$  (cf. note 20), it might be best to claim that *kaoi* is adjoined to the leftmost periphery of its clause. Cf. Brandon and Seki (1981) for evidence of this type of phenomenon in other Amazon languages.

23. While it might be more accurate to label the second position clitics as **INFL-clitics**, we will not pursue this here.

## References

- Aoun, J. 1979. A short note on cliticization. Mimeographed, MIT.
- Bok-Bennema, R. 1981. Clitics and Binding in Spanish. Levels of syntactic representation, ed. by R. May and J. Koster. Dordrecht: Foris.
- Borer, H. 1981. Parametric variation in clitic configurations. Cambridge: MIT dissertation.
- Brandon, F.R. and L. Seki. 1981. A note on COMP as a universal. LI 12,659-665.
- Chomsky, N. 1965. Aspects of the theory of Syntax. Cambridge: MIT Press.
- . 1971. Deep structure, surface structure, and semantic interpretation. Semantics, ed. by D. Steinberg and L.A. Jakobovits. London: Cambridge University Press.
- . 1977. On Wh-movement. Formal syntax, ed. by P. Culicover et al. New York: Academic Press.
- . 1980. On binding. LI 11,1-46.
- . 1981. Lectures on government and binding. Dordrecht: Foris.
- . 1982. Some concepts and consequences of the theory of government and binding. Cambridge: MIT Press.
- . and M. Halle. 1968. The sound pattern of English. New York: Harper and Row.
- Everett, D. 1981. Algumas implicações epistemológicas do desenvolvimento da teoria linguística. A paper presented at the VI Encontro Nacional de Linguística, PUC-RJ, Rio de Janeiro.
- . 1983a. A língua Pirahã e a teoria da sintaxe. Campinas, Brazil: Universidade Estadual de Campinas. Dissertation.
- . 1983b. Sobre a interpretação da referencia na teoria de regencia e vinculação. A paper presented at the VIII Encontro Nacional de Linguística, PUC-RJ, Rio de Janeiro.

- . MS. Pirahã. To appear in Handbook of Amazonian languages, ed. by D. Derbyshire and G. Pullum.  
Fieigo, R. 1980. Surface Structure. Cambridge: Harvard University Press
- Givón, T. 1976 Topic, pronoun, and grammatical agreement. Subject and topic, ed. by C. Li. New York: Academic Press.
- Greenberg, J. 1966. Some universals of grammar with particular reference to the order of meaningful elements. Universals of language, ed. by J. Greenberg. Cambridge: MIT Press.
- Hale, K. 1983. Warlpiri and the grammar of non-configurational languages. Natural Language and Linguistic Theory 1,5-47.
- Higgenbotham, J. 1983. Logical form, binding, and nominals. LI 14.395-520.
- Huang, J. 1981. Move WH in a language without WH movement. Linguistic Review 1,369-416.
- . 1983. A note on binding theory. LI 14.554-561  
v.d. Hulst, H. and N. Smith. 1982. The structure of phonological representations. Vol 1 1 2. Dordrecht: Foris.
- Jaeggli, O. 1980. On some phonologically null elements in syntax. MIT dissertation.
- . 1982. Topics in romance syntax. Dordrecht: Foris
- Kaisse, E.M. 1982. Sentential clitics and Wackernagel's law. Proceedings of the first West Coast conference on formal linguistics, ed by D.P. Flickinger et al. Stanford: Stanford Linguistics Association.
- Kayne, R. 1975. French syntax. Cambridge: MIT Press.
- . 1981. Unambiguous paths. Levels of syntactic representation, ed. R. May and J. Koster. Dordrecht: Foris.
- May, R. 1977. The grammar of quantification. Cambridge: MIT dissertation.
- Newmeyer, F. 1980. Linguistic theory in America. New York: Academic Press.
- Reinhart, T. 1982. Pragmatics and linguistics: an analysis of sentence topics. Bloomington: Indiana University Linguistics Club.

- Reuland, E. 1983. Governing -ing. *LI* 14.101-146.
- v. Riemsdijk, H. and E. Williams. 1981. NP-structure. *The Linguistic Review* 1.171-217.
- Safir, K. 1981. Inflection - government and inversion. *The Linguistic Review* 1.369-416.
- . 1982. Syntactic chains and the definiteness effect. MIT dissertation.
- Shaul, D.L. 1983. Clisis and word order in Tepiman. *Interplay*, ed. by J.F. Richardson et al. Chicago: CLS.
- Steele, S. et al. 1981. *An encyclopedia of aux.* Cambridge: MIT Press.
- Stowell, T. 1981. The origins of phrase structure. MIT dissertation.
- . 1982. Subjects across categories. *The Linguistic Review* 2.285-312.