

SOME SYNTACTIC RULES IN BARON LONG DIEGUEÑO*

1. Ross (1967) tentatively proposes an output condition on the acceptability of grammatical English sentences, and he suggests that this condition may turn out to be a linguistic universal.¹ The condition is as follows:

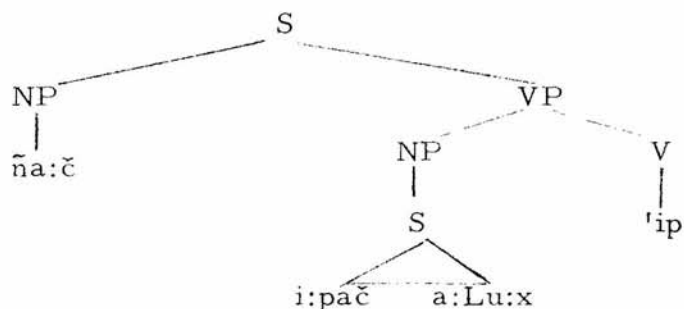
Grammatical sentences containing an internal NP which exhaustively dominates an S are unacceptable.

Data from the Baron Long dialect of Diegueño show that this condition does not apply to this language. The usual form of sentences containing a complement clause is one in which the clause is embedded between the subject and verb of the higher S. If we adopt Ross' own notion of the deep structure of complement clauses² (which seems plausible to me), then (1) will have a surface structure identical in all relevant respects to its deep structure, given in (2):

- (1) ña:č i:pač a:Lu:x 'ip
 I man snore hear³

I hear the man snoring

(2)



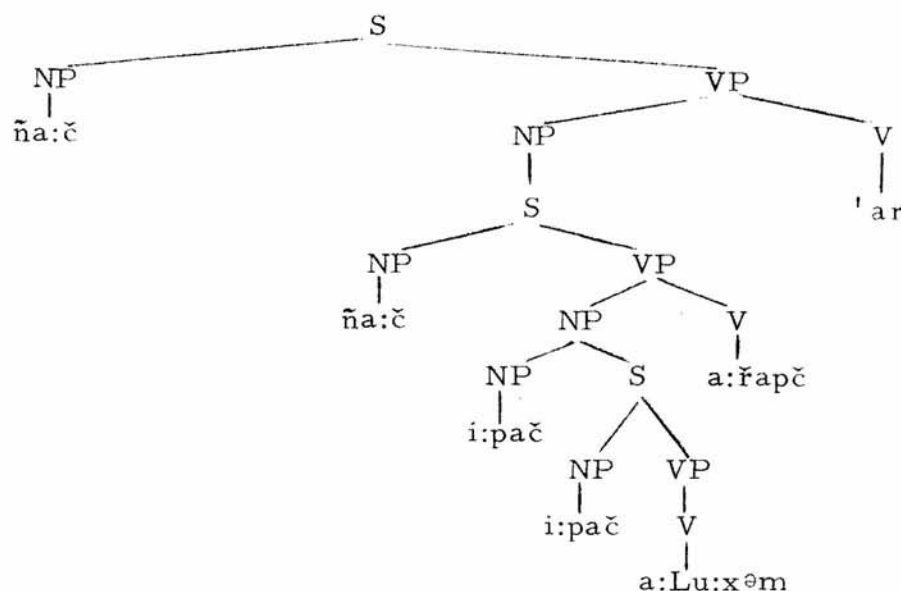
More complicated sentences can also appear with completely embedded clauses:

- (3) ña:č i:pač a:Lu:xəm a:řapč 'ar
 I man snore hit want

I want to hit the man who is snoring

As with (1), the deep and surface structures of (3) are essentially identical:

(4)

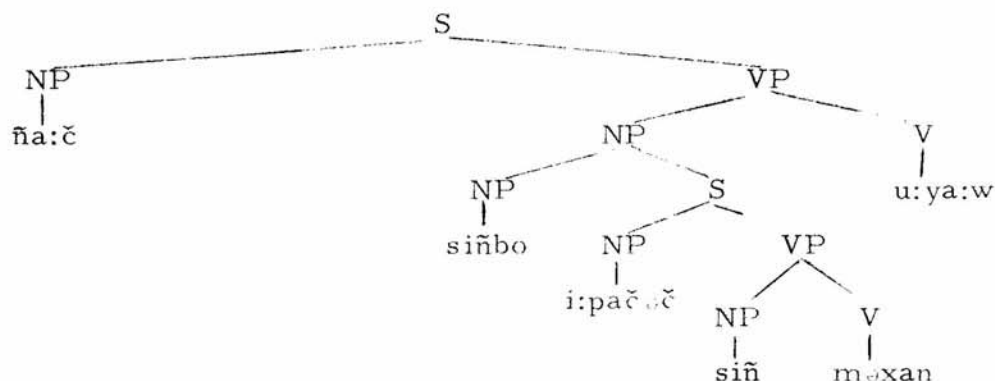


Comparison of (4) with (3) shows that Baron Long must have a rule that deletes one of two NP's that are, in some sense, identical. That is, in order to turn the deep structure (4) into the surface structure of (3) one of the occurrences of ñā:č and one of the occurrences of i:pač must be deleted. How is this rule to be formulated? We might infer from (4) that one of two identical NP's is deleted when they are contiguous. However, sentences (5) and (7) show that this formulation is inadequate. Sentence (5) has the deep structure (6), and sentence (7) the deep structure (8):

- (5) ñā:č siñbo u:ya:w i:pačəč məxan
 I lady know man like

I know the lady who the man likes

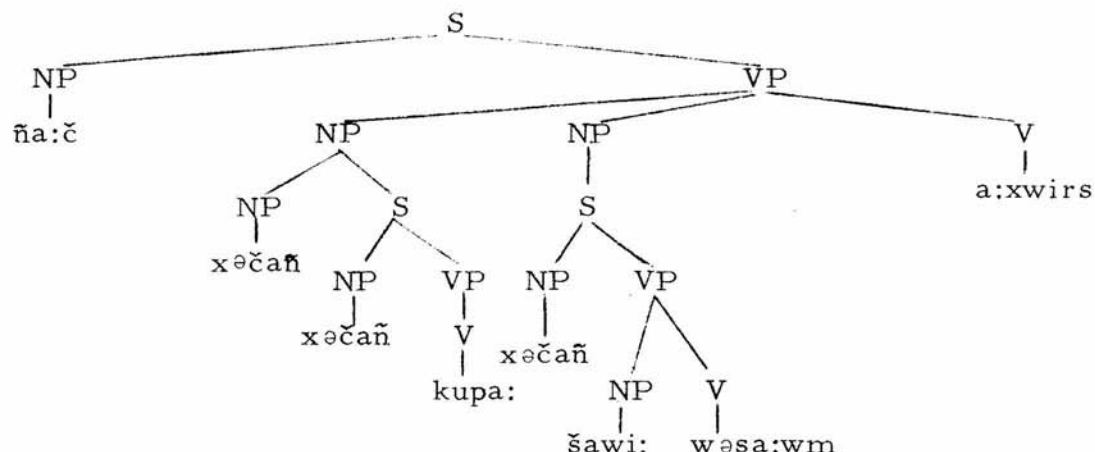
(6)



- (7) ña:č xəčaň kupa: šawi: wəsa:wm a:xwirs
 I girl come mush eat force

I forced the girl who came to eat the mush

(8)



Sentences (5) and (7) have both undergone rules that we have not yet discussed-- sentence (5) a rule that postposed the clause i:pačəč məxan, and sentence (7) a nominalization rule. These facts are, however, irrelevant to the formulation of the rule that deletes a NP under identity. In (6) the occurrence of siň in the embedded clause is not contiguous to the occurrence of the identical NP in the higher clause. Similarly in (8), the occurrence of xəčaň in the clause xəčaň šawi: wəsa:wm is not contiguous to the occurrence of xəčaň in the highest S. This leads to the formulation (9), using the device of indices to indicate the required kind of identity:

(9) Identical-NP-Deletion

X	NP _i	Y	NP _i	Z
1	2	3	4	5 \Rightarrow OBLIG.
1	2	3	∅	5

condition: 2 commands 4⁴

We will see later why it is the second of the two identical NP's that must be deleted by this rule. The rule must be obligatory because of the ungrammaticality of sentences like (10) and (11), where repeated occurrences of the same word are taken to be identical in the relevant sense:

- (10) *ña:č siňbo u:ya:w i:pačəč siňbo məxan
 I lady know man lady like

- (11) *na:č siñ a:xwirs siñ šawi: wəsa:w
 I lady force lady mush eat

Let us now consider sentences with subordinate clauses not entirely embedded--that is, sentences that have undergone the post-posing rule referred to with respect to (5). For example:

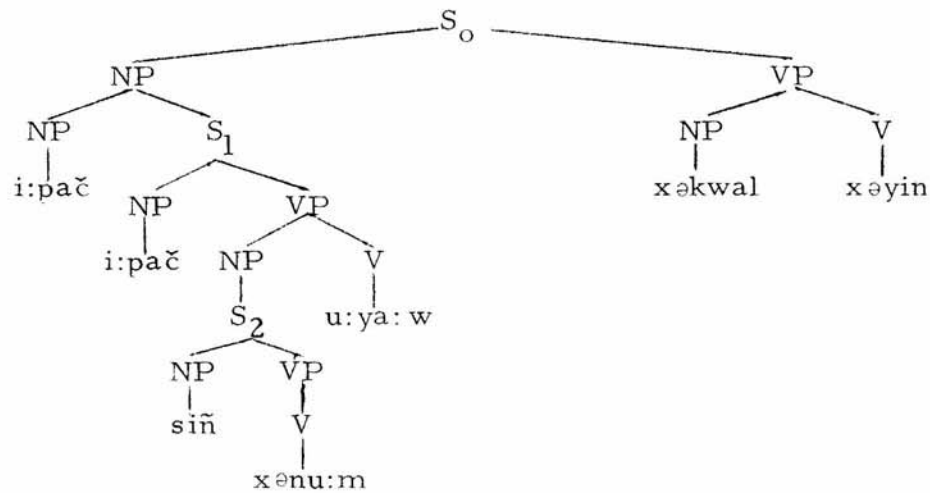
- (12) i:pač xəkwal xəyin siñ xənu:m u:ya:wč
 man child carry lady sick know

This sentence has two readings:

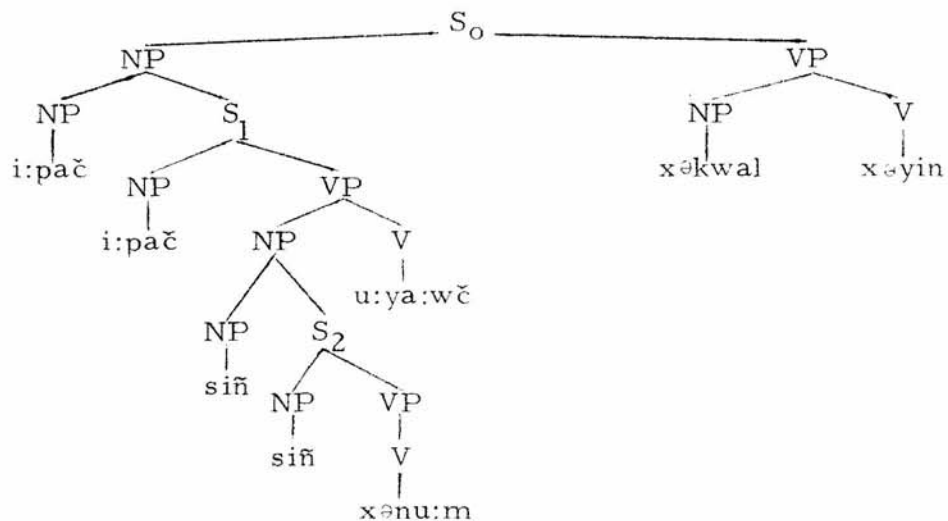
- (a) The man who knows that the lady is sick is carrying her child.
 (b) The man who knows the lady who is sick is carrying her child.

The two readings (12a) and (12b) would have the deep structures (13a) and (13b), respectively:

(13a)



(13b)



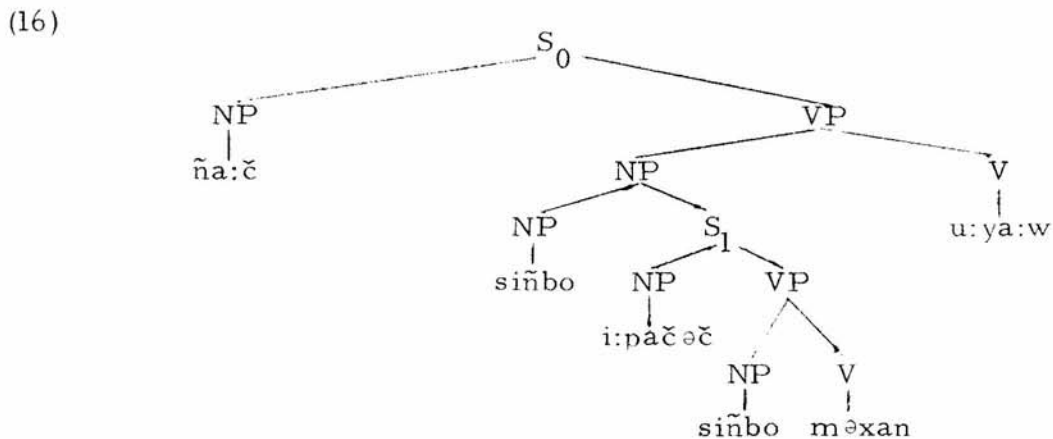
From the evidence of (12a,b) and (13a,b) we could conclude that the rule in question simply moves an embedded VP to the end of the higher S. That is, in order to derive (12a) from (13a), first delete the rightmost occurrence of i:pač by the rule of Identical-NP-Deletion, and then move the embedded $VP \left[NP \left[S \left[\underline{si\tilde{n} \ x\acute{a}nu:m} \right] S \right] NP \ V \left[u:ya:w\check{c} \right] V \right] VP$ to the end of the higher S. Similarly, to derive (12b) from (13b), first delete the rightmost occurrences of both i:pač and siñ, and then move the embedded $VP \left[NP \left[NP \left[\underline{si\tilde{n}} \right] NP \ S \left[\underline{si\tilde{n} \ x\acute{a}nu:m} \right] S \right] NP \ V \left[u:ya:w\check{c} \right] V \right] VP$ to the end of the higher S.

However, it is possible to show that the rule in question moves not simply an embedded VP, but an entire S. Consider the following examples:

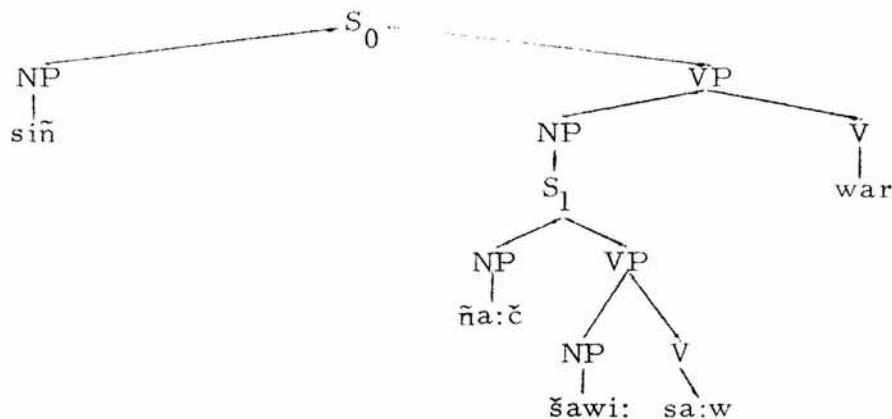
- (14) $\tilde{n}a:\check{c}$ $si\tilde{n}bo$ $u:ya:w$ $i:pa\check{c}\acute{e}\check{c}$ $m\acute{e}xan$ (same as (5))
 I lady know man like
 I know the lady who the man likes.

- (15) $si\tilde{n}$ war $\tilde{n}a:\check{c}$ $\check{s}awi:$ $sa:w$
 lady want I mush eat
 The lady wants me to eat the mush.

The deep structures of (14) and (15) are (16) and (17), respectively:



(17)



In (16), Identical-NP-Deletion will delete the rightmost occurrence of siñbo, but S_1 continues to branch and will not be pruned.⁵ As (14) shows, S_1 is moved in its entirety to the end of the sentence. In (17), Identical-NP-Deletion does not apply (since there are no identical NP's), and S_1 is moved as is to the end of the sentence. The NP that directly dominates S_1 will now dominate nothing and will presumably be pruned. Let us term the rule in question S-Movement and state it as follows:

- (18) S-Movement⁶
- $$S \left[\begin{array}{ccc} X & S & Y \end{array} \right]_S$$
- $$\begin{array}{ccc} 1 & 2 & 3 \end{array} \Rightarrow \text{OPT.}$$
- $$\begin{array}{ccc} 1 & \emptyset & 3 + 2 \end{array}$$

2. Let us now consider the nominalization rule referred to with respect to (7). This rule, which Langdon terms "KW-Nominalization", has the effect of adjoining to the verb a prefix, which has the alternate pronunciations kw and ku. Some examples of sentences that have undergone this rule follow:

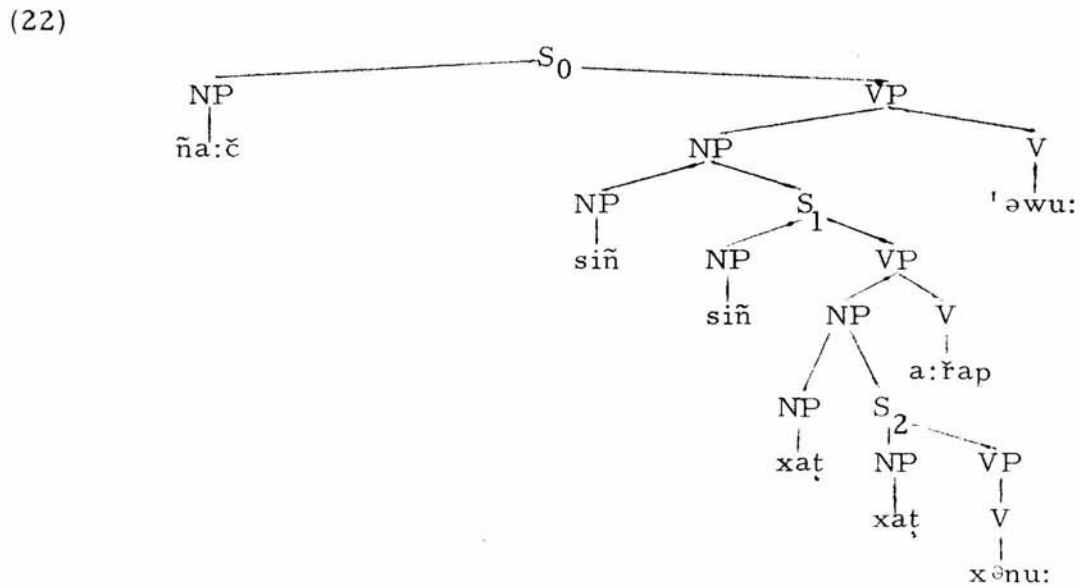
- (19) ña:č kwəxwaɬ čəmxan
 I red-one like
I like the red one.

- (20) kəyu: xa: kwənu: kəwu:
 come water that-runs see
Come see the water that runs. (i. e., the river)

- (21) ña:č siñ xaṭ kwəxnučəč kwa:řap 'əwu:
 I lady dog that-is-sick who-hit see

I saw the lady who hits the sickly dog.

Sentences (19)-(21) have verbs with prefixed KW functioning as the verbs of relative clauses. For example, (21) contains two relative clauses, and the verbs of both of these clauses have a prefixed KW. The deep structure of (21), with relative clauses S_1 and S_2 , is given in (22):



Now consider examples (23) and (24), which show that KW-Nominalization can, in fact, only apply to relative clauses and not to complement clauses:

- (23) *ña:č siñ kupa: a:xwirs
 I lady who-came force
 (cannot mean, "I forced the lady to come.")

- (24) ña:č siñ kupa: u:ya:w
 I lady who-came know
 I know the lady who came.
 (but cannot mean, "I know that the lady came.")

Nor can sentences that have undergone KW-Nominalization be moved by the rule S-Movement:

- (25) ña:č siñ xaṭ xənu: a:řap 'əwu:
 I lady dog sick hit see
 (a) I saw the lady hit the sick dog.
 (b) I saw the lady who hit the sick dog.

- (26) ña:č siñ 'əwu: xaɬ xənu: a:řap
 I lady see dog sick hit

I saw the lady who hit the sick dog.

- (27) *ña:č siñ wu: xaɬ kwəxnučəč kwa:řap
 I lady see dog that-is-sick who-hit

(cannot mean the same as (21).)

Sentence (25) on the reading (25b) has not undergone KW-Nominalization and can therefore undergo S-Movement, giving (26). But (21) has undergone KW-Nominalization, and (27) cannot be derived from (21) by S-Movement.

We can now state the rule of KW-Nominalization, but first let us take note of a general restriction on the rule, which I will give here informally, rather than trying to build it into the statement of the rule itself. KW-Nominalization can apply only to third-person sentences where the verb has no syntactic affixes.

- (28) KW-Nominalization

$$\begin{array}{ccccccc}
 \text{NP} & \left[\text{NP} & \text{S} & \left[\text{X} & \text{V} & \text{Y} \right] & \text{S} \right] & \text{NP} \\
 & 1 & & 2 & 3 & 4 & & \Rightarrow \text{OPT.} \\
 & 1 & & 2 & \text{KW} + 3 & 4 & &
 \end{array}$$

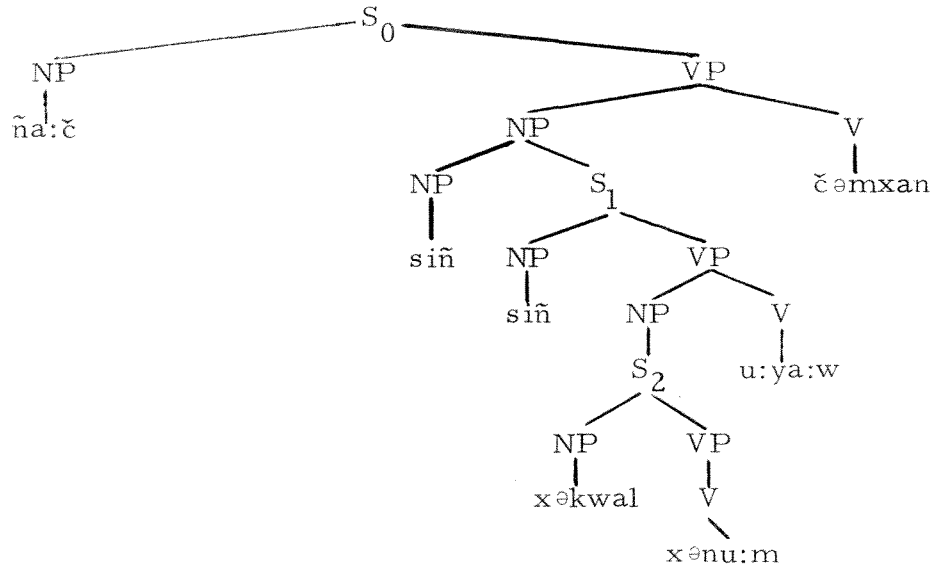
Now note that we must put another restriction on the rule--namely, that the V of term 3 must be the main verb of the bracketing S. That this restriction on (28) is necessary is shown by sentence (29), with deep structure (30):

- (29) ña:č siñ xəkwal xənu:m u:ya:w čəmxan
 I lady child sick know like

I like the lady who knows that the child is sick.

(This sentence is multiply ambiguous, but the above reading is the relevant one for this discussion.)

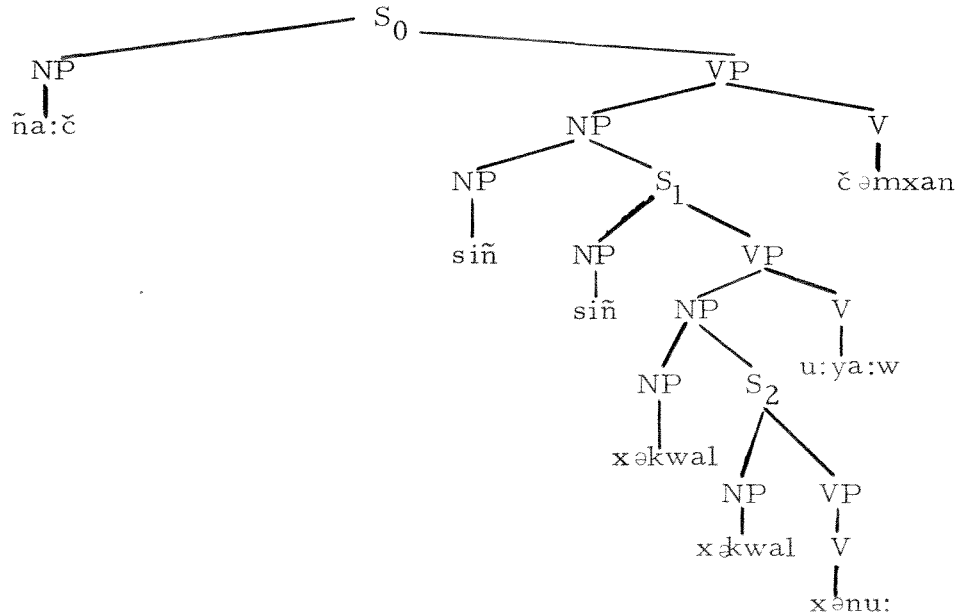
(30)



On the S_0 cycle, the V in S_2 , $x\acute{e}nu:m$, could fulfill the structural description of (28), deriving (31), unless the condition in question were in effect. Note that, although (31) is a grammatical sentence, it is not synonymous with (29) and does not have the deep structure (30). Rather it has the deep structure (32):

- (31) $\tilde{n}a:\tilde{c}$ $si\tilde{n}$ $x\acute{e}kw\acute{a}l$ $kw\acute{e}xnu\check{c}\acute{e}\check{c}$ $u:ya:w$ $\check{c}\acute{a}mxan$
 I lady child who-is-sick know like
 I like the lady who knows the sickly child.

(32)



Thus the necessity for the condition on term 3 of (37) is established.

3. We can now examine the operation of a fourth rule of Baron Long, which has not yet been mentioned. This rule is of a type that has been termed "Scrambling".⁷ That is, the effect of the rule is, under certain conditions, to more-or-less freely rearrange the constituents of sentences. The following sentences illustrate some of the permitted and nonpermitted consequences of the operation of this rule:

- (33) xatbo siñəč a:řap
 dog girl hit

The girl hit the dog.

- (34) ña:č čəmxan siñbo
 I like lady

I like the lady.

- (35) šawi:bo wəsa:w i:pačəč
 mush eat man

The man ate the mush.

- (36) wəwu: i:pačəč siñbo
 see man lady

The man saw the lady.

- (37) čəmxan siñbo i:pačəč
 like lady man

The man likes the lady.

- (38) *šawi: wəsa:w siñ war
 mush eat lady want

(cannot mean, "The lady wants to eat the mush.")

- (39) *i:pač xakwal wəpis ña:č 'u:ya:w⁸
 man child kiss I know

(cannot mean, "I know the man who kissed the child.")

- (40) *ña:č i:pačbo siñəč wəpis 'u:ya:w⁹
 I man lady kiss know

(cannot mean, "I know that the lady kissed the man.")

Sentences (33)-(40) show that in simple sentences the subject, object, and verb can appear in any of the six possible orders:

SOV (all previous examples)

OSV (33)

SVO (34)

OVS (35)

VSO (36)

VOS (37)

Sentence (38) shows that when the object of the verb is a sentence it cannot be preposed. Sentence (39) shows that when the object of the verb is an NP with a relative clause, it cannot be preposed. Sentence (40) illustrates the impossibility of rearranging the constituents of simple sentences when they are embedded in other sentences. Note that "rearrange" is taken to mean any alteration from the SOV order, which we have assumed to be basic (for discussion of this point see Section 4). From these examples we can conclude that Scrambling in Baron Long can apply only to simple sentences, that is, sentences that do not have any other sentences embedded in them. We can formulate the rule as follows:

(41) Scrambling

$$\begin{array}{c}
 S \left[\begin{array}{ccc} \text{NP} & \text{NP} & \text{V} \end{array} \right] S \\
 \begin{array}{ccc} 1 & 2 & 3 \end{array} \Longrightarrow \text{OPT.} \\
 \nrightarrow (1 \quad 2 \quad 3)
 \end{array}$$

condition: neither term 1 nor term 2 dominates S.

I have used the symbol \nrightarrow as a notational device to indicate that the structural change of (41) is any of the six possible linear orders of terms 1, 2, and 3.

4. Baron Long is a language that contains a rule of Gapping, in Ross's sense of the term.¹⁰ The permitted and nonpermitted interactions of the rules Gapping and Scrambling have been used by Ross in determining the underlying, or basic, order of constituents in languages.¹¹ Consider a paradigm of sentences involving the rules Gapping and Scrambling, like the paradigms adduced by Ross for other languages:

The following sentences are all intended to mean, "The man hit the dog, and the lady the cat."

i:pačəč = man + subject marker
 siñəč = lady + subject marker
 xaɬbo = dog + object marker
 ga:tbo = cat + object marker
 a:řap = hit

- (42) i:pačəč a:řap xaɬbo, siñəč ga:tbo
 (43) *i:pačəč xaɬbo, siñəč a:řap ga:tbo
 (44) i:pačəč xaɬbo a:řap, siñəč ga:tbo
 (45) i:pačəč xaɬbo, siñəč ga:tbo a:řap
 (46) xaɬbo a:řap i:pačəč, ga:tbo siñəč
 (47) *xaɬbo i:pačəč, ga:tbo a:řap siñəč
 (48) xaɬbo i:pačəč a:řap, ga:tbo siñəč
 (49) xaɬbo i:pačəč, ga:tbo siñəč a:řap
 (50) a:řap i:pačəč xaɬbo, siñəč ga:tbo
 (51) *i:pačəč xaɬbo, a:řap siñəč ga:tbo
 (52) a:řap xaɬbo i:pačəč, ga:tbo siñəč
 (53) *xaɬbo i:pačəč, a:řap ga:tbo siñəč

From (42)-(53), we can see that the following sequences are not permitted:

- *SO, SVO (43)
 *OS, OVS (47)
 *SO, VSO (51)
 *OS, VOS (53)

Following Ross, let us assume that Gapping is an "anywhere" rule (that is, that it can apply at any point in a derivation). This means that Gapping could apply before and/or after Scrambling. By an argument exactly parallel to the argument that Ross gives for Russian, we could show that it is only by assuming an underlying SVO order that we could derive just the paradigm of sentences (42)-(53). I will not rephrase Ross' argument here, however, because further considerations show beyond doubt that

Baron Long cannot have a basic SVO order and that the basic order must be, as we have assumed all along, SOV. It is clear that this casts doubt on the validity of the argument from Gapping for determining the underlying order of constituents in other languages.

First, there are two considerations with respect to what we feel to be the natural or typical state of affairs in language: (1) if a language, like Baron Long, has subject and object markers that are obligatory in all instances except if subject, object, and verb are in a particular linear order, in which case they can optionally be deleted, then we would like to say that the latter "unmarked" order is somehow basic in the language; (2) if, for no other apparent reason, the overwhelming majority of sentences produced by native speakers of a language are in a particular order, then again we would like to say that that order is somehow basic. Both (1) and (2) are true of SOV order in Baron Long. (See also Elgin (1969), in this volume.)

Second, and much more important, it is clear that assuming an underlying SVO order for Baron Long would make it necessary to put a very strange and totally ad-hoc constraint on the rule Scrambling. For only SOV and never SVO order is possible in embedded sentences. Therefore, the Scrambling rule would have to be allowed to apply to embedded sentences, but in a very restricted way, with only the SOV output permitted. However, on simple sentences Scrambling would have to be allowed to apply freely. Furthermore, the application of Scrambling in embedded clauses would have to be made obligatory, whereas it is optional in simple sentences. These facts are evidence of a very strong kind that the underlying order of constituents in Baron Long cannot be SVO and must be SOV. As an alternative to putting the above restrictions on Scrambling one could postulate a rule that obligatorily applies to embedded clauses and that partially duplicates Scrambling, in that the rule would take as input SVO orders and give as output SOV orders. But this solution, being totally ad hoc, is hardly attractive.

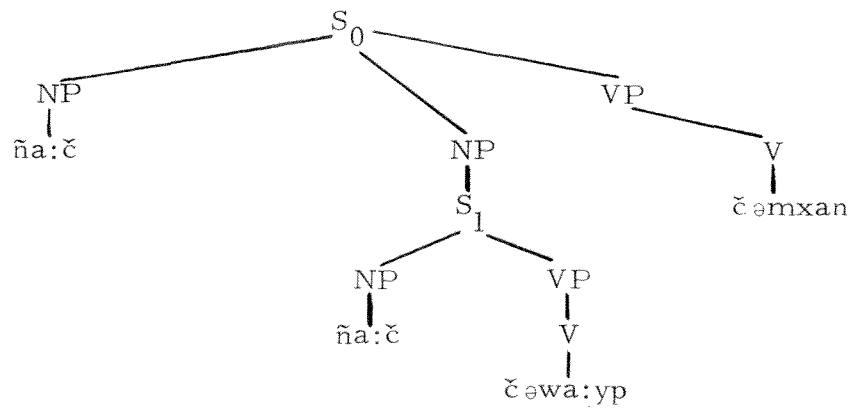
5. In this section we will point out a further constraint on Baron Long sentences and we will show that this constraint is best formulated as an output condition--that is, **as** a condition on well-formed surface structures. Consider the following sentences and their associated deep structures:

(54) (a) $\tilde{n}a:\tilde{c}$ $\tilde{c}əwa:yp$ $\tilde{c}əm xan$
 I talk like

 (b) $*\tilde{n}a:\tilde{c}$ $\tilde{c}əm xan$ $\tilde{c}əwa:yp$
 I like talk

I like to talk.

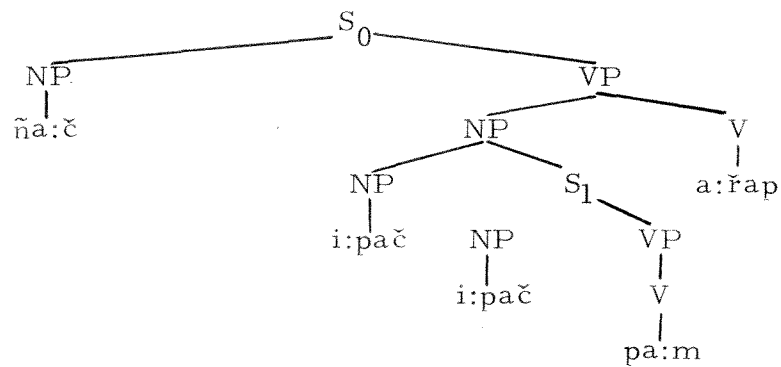
(55)



- (56) (a) ña:č i:pač pa:m 'a:řap
 I man come hit
- (b) *ña:č i:pač 'a:řap pa:m
 I man hit come

I hit the man who came.

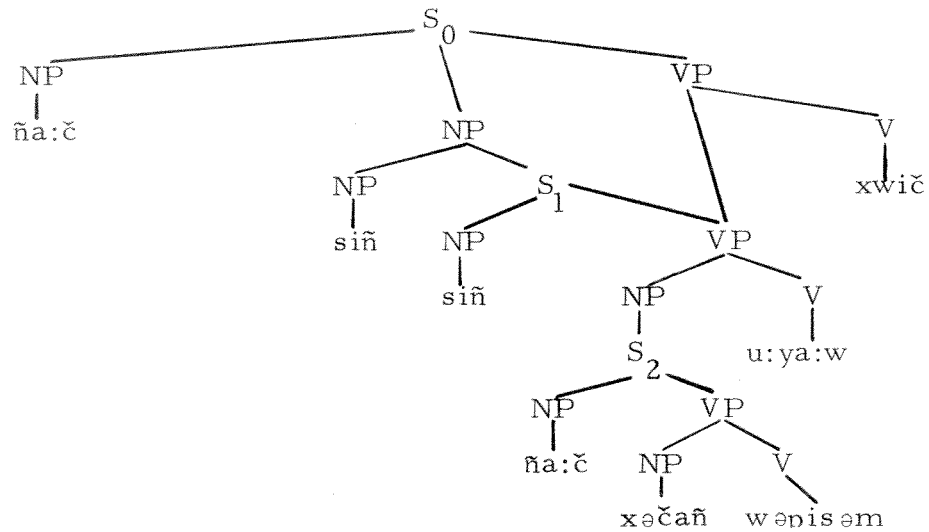
(57)



- (58) (a) ña:č siñ xwič xəčaň wəpisəm u:ya:w
 I lady hate girl kiss know
- (b) *ña:č siñ xwič u:ya:w xəčaň wəpisəm
 I lady hate know girl kiss

I hate the lady who knows that I kissed the girl.

(59)



The three ungrammatical sentences (54), (56), and (58) have all been produced by the application of S-Movement to embedded sentences S_1 , and, in the case of (58), to S_2 . In order to derive (58) from (59), S-Movement applies, on the S_1 cycle, moving S_2 to the end of S_1 . Then, on the S_0 cycle, S-Movement applies again, moving S_1 to the end of S_0 , and Identical-NP-Deletion applies, erasing the rightmost occurrences of ñā:č and siñ. The derivations of (54) and (56) from (55) and (57), respectively, are obvious and straightforward. What is the condition, then, that disallows these sentences? Clearly it is not, as (54) and (56) seem to suggest, that S-Movement cannot apply to sentences with intransitive verbs. For in (58) none of the verbs is intransitive, and it would be satisfying to disallow all three sentences for the same reason. An examination of the surface word orders of the three sentences gives the clue. In each case there is a lower verb immediately following a higher verb: in (54), čəwa:yp immediately follows čəmɣan; in (56), pa:m immediately follows a:řap; and in (58), u:ya:w immediately follows xwič. Thus, we can state the following condition:

- (60) In Baron Long surface structures there may not exist a sequence V_1V_2 such that V_2 is dominated by S_i and V_1 is dominated by S_j and S_j dominates S_i .

The fact that it is possible to rule out all of (54), (56), and (58) for the same reason, and the fact that it is impossible to rule out (58) by a condition on the structural description of S-Movement, provide strong motivation for condition (60).

FOOTNOTES

* Diegueño, a member of the Yuman language family, comprises a number of dialects spoken in and around San Diego County. The dialect under consideration in this paper is that spoken on the Baron Long reservation near Alpine, California. Henceforth I will refer to this dialect simply as "Baron Long".

1. The condition is given as example (3.27) on p. 57. The suggestion that, in a modified form, it may be a linguistic universal is on p. 69.
2. There is much controversy today over the form and the content of the phrase structure rules of grammars. In this paper I will adopt underlying structures like those proposed in recent work by Lakoff and Ross. Also, for evidence that in Diegueño, as in English (see Rosenbaum, 1967), complement sentences are NP's, see Langdon (1966), sec. 842.2, p. 251.
3. The word-for-word translations omit translation of many grammatical morphemes. These morphemes are described in Langdon (1966), as well as in Elgin (1969), in this volume.
4. For definition of the notion of 'command' see Langacker (1969).
5. See Ross (1966) for discussion of node-pruning and the conditions under which it operates.
6. There is a restriction on term 3 of (18). Namely, the variable Y cannot dominate an S. In other words, the S being moved cannot be moved around another S. This is shown by examples like the following:

- (1) *i:pač xəkwal xənu:m u:ya:w siñ wəpɪs
 man child sick know lady kiss
 (cannot mean, "The man who kissed the lady knows that the child is sick.")
- (2) *ña:č i:pač kwinkuy pa:m kuna:p šawi: wəsa:wm
 I man lady come tell mush eat
 (cannot mean, "I told the man who ate the mush that the lady came.")

In (1), the relative clause siñ wəpɪs has been moved around the complement clause xəkwal xənu:m. In (2), the relative clause šawi: wəsa:wm has been moved around the complement clause kwinkuy pa:m. Note that the ungrammaticality of (1) and (2) cannot be explained by saying that the movement creates an uncertainty about which NP the relative clause modifies, since (3) is ungrammatical for the same reason, yet the relative clause šawi: wəsa:w cannot apply to any NP other than i:pač.

- (3) *ña:č i:pač wa: wəLap kuna:ɒ šawi: wəsa:wm
I man house burn tell mush eat

(cannot mean, "I told the man who ate the mush that the house burned down.")

This restriction on term 3 follows from a principle put forth in Ross (1967a). On p.307, Ross states in example (5.58):

Any rule whose structural index is of the form . . . A Y,
and whose structural change specifies that A is to be adjoined
to the right of Y, is upward bounded.

"Upward bounded" means that A cannot be moved beyond the first sentence up. It is clear that S-Movement as formulated in (18) is the type of rule referred to in Ross' condition. Thus, the restriction on term 3 of (18) does not have to be stated as part of the rule of S-Movement; rather, it follows from a principle of universal grammar.

A further comment on S-Movement is appropriate here. Note that this rule can perform the same operation as the rule of English Extrapolation from NP, discussed in Ross (1967a). In addition, the Baron Long rule can also postpose complement clauses. This suggests that S-Movement might be a partially universal rule, having the two functions of extraposing relative clauses and (in languages where the object precedes the verb) of extraposing complement clauses.

7. For discussion of rules of this type see Ross (1967a), pp. 74-80.

8. Sentence (39) could be given as a translation of "I know that the man kissed the child". However, it seems impossible to decide whether, on this reading, (39) should be thought of as one sentence or two. That is, on this reading, it is just as plausible to translate (39) as the sequence of sentences: "The man kissed the child." "I know it." Since considering sentences like this to be single sentences would greatly limit the generality of the restrictions on Scrambling, I will assume the 2-sentence analysis.

9. Sentence (40) is only ungrammatical on the reading indicated. It is perfectly grammatical meaning, "I know the man who the lady kissed."

10. Ross, (1967b).

11. Ibid.

BIBLIOGRAPHY

- Chomsky, N., Aspects of the Theory of Syntax, MIT Press: Cambridge, 1965.
- Elgin, Suzette, "Some Remarks on Disambiguating Mechanisms in Diegueño", 1969 (in this volume).
- Katz, J. J. and Postal, P., An Integrated Theory of Linguistic Descriptions, MIT Press: Cambridge, 1964.
- Lakoff, G., On the Nature of Syntactic Irregularity, Report No. NSF-16, Cambridge, 1965.
- _____, "Deep and surface grammar". Unpublished MS, 1966.
- Langdon, M., A Grammar of Diegueño: the Mesa Grande Dialect. Unpublished doctoral dissertation, University of California, Berkeley, 1966.
- Rosenbaum, P. S., The Grammar of English Predicate Complement Constructions, MIT Press: Cambridge, 1967.
- _____, English Grammar II, IBM Research Monograph RC 2070, Yorktown Heights, New York, 1968.
- Ross, J. R., "A proposed rule of tree-pruning", Report No. NSF-17, Cambridge, 1966.
- _____, Constraints on Variables in Syntax. Unpublished doctoral dissertation, MIT, Cambridge, 1967a.
- _____, "Gapping and the order of constituents." Unpublished MS, 1967b.