

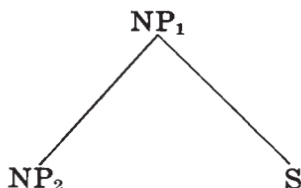
ON THE COMPLEX NP IN PORTUGUESE

Marco A. de Oliveira

Introduction

Ross (1968) points out that no constituent can be reordered outside an embedded sentence if this sentence is preceded by a lexical item, being both the lexical item, and the embedded sentence dominated by the same NP. Ross points out, too, that we have to distinguish between "lexical items", on the one hand, and the abstract pronoun *it*, on the other hand, since in a structure like

(1)



a constituent can be transformationally reordered out of the S node if NP_2 dominates *it*, but not if NP_2 dominates a "lexical item" Compare the sentences (2) - (4) :

- (2) (a) I saw the man who wrote this book.
(b) *This is the book that I saw the man who wrote.
- (3) (a) I believe the claim that Mary wrote this book.
(b) *This is the book that I believe the claim that Mary wrote.
- (4) (a) I'm going to see to it that my wife buys this hat.
(b) This is the hat which I'm going to see to it that my wife buys.

Ross proposes, then, that the theory of grammar must incorporate the following constraint:

(5) *The Complex NP Constraint*

“No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that phrase by a transformation”.

Koutsoudas (1973), assuming that extrinsic order is not necessary for the explanation of facts about natural languages, provides evidence that the *Complex NP Constraint* must be a derivational constraint and not a local constraint. Koutsoudas considers cases similar to (6) - (7)

- (6) (a) The man who bought the car jumped over the fence.
(b) *The car, the man who bought jumped over the fence.
- (7) (a) The man jumped over the fence who bought the car.
(b) *The car, the man jumped over the fence who bought.

The ungrammaticality of (6b) is predictable from (5). The topicalized element, *the car*, was moved out of the complex NP. But in (7) the relative clause *who bought the car* was extraposed from the complex NP. So, we do not have a complex NP anymore and the ungrammaticality of (7b) is not predictable from (5). If extrinsic ordering is not part of grammar, the only way to account for (7b) is to make the *Complex NP Constraint* a derivational constraint instead of a local one. This new version of the *Complex NP Constraint* is stated by Koutsoudas as

(8) *The Complex NP Constraint (revised)*

“If in a line in a derivation there is a complex noun phrase, NP₁, consisting of a lexical head NP₂ and a sentence, S₁, then there cannot be another line in this derivation in which an element which was contained in S₁ is not dominated either by NP, or by S₁”.

Kuno (1976) proposes a different explanation for part of these facts, in terms of his *Thematic Constraint on Relative Clauses*, restated here as (9), to account for the ungrammaticality of sentences like (10):

(9) *The Thematic Constraint on Relative Clauses*

“A relative clause must be a statement about its head noun.”

(10) *This is the child who John married a girl who dislikes.

Kuno's argument, as far as I understood it, is as follows. In a sentence like (10) :

(11) John married a girl who dislikes this child
the relative clause *who dislikes this child* is 'a statement about' a girl. So, a girl in (11) is the theme of the relative clause. On the other hand, it is not possible to interpret *this child* in (11) as a theme. Therefore, *this child* cannot be the theme in (10), and that is why (10) is ungrammatical. Kuno also says that in a sentence like

(12) There is nobody who is willing to adopt this child

it is easy to interpret *this child* as the theme of the sentence. On the other hand the relative clause *who is willing to adopt this child* is 'a statement about' nobody, which is, according to Kuno, "semantically transparent". Therefore, a sentence like

(13) This is the child who there is nobody who is willing to adopt

is a possible one, according to Kuno. Notice that a sentence like (13) would be blocked by both (5) and (8). Notice also that a sentence like (3b) would have to be equally possible in Kuno's framework since *the claim* is not the theme and so there is no reason why *the book* could not be extracted out of the complex NP. Unfortunately Kuno omits his opinion about cases like (3b), which are correctly blocked by (5) and (8).

1. *The Complex NP in Portuguese*

The *Complex NP Constraint*, as noted before, was designed to incorporate two different kinds of sentences: relative clauses and *fact-S clauses*. In both cases the syntactic configuration of (1) holds. Consider the following sentences of Portuguese:

(14) João viu o menino que comprou o camelo.

'João saw the boy who bought the camel'.

The NP *o camelo* 'the camel' in the relative clause cannot be moved out of the complex NP, as shown in (15) - (17)

(15) *Foi um camelo que João viu o menino que comprou.

'It was a camel that João saw the boy who bought'.

(16) *Este é o camelo que João viu o menino que comprou.

'This is the camel that João saw the boy who bought'.

(17) *O que João viu o menino que comprou?

'What did João see the boy who bought?

Both (5) and (8) rule out sentences (15) - (17). Constraint (9) provides an explanation why (15) - (17) are ungrammatical. Since *o camelo* cannot be interpreted as the theme in (14), any syntactic transformation which would make it a theme would violate (9). Consider now

(18) Não há ninguém que queira adotar esta criança.

'There is nobody who is willing to adopt this child'.

Since *ninguém* 'nobody' is a 'semantically transparent' theme it is always possible to consider *esta criança* 'this child' as the theme in (18). But, if we move *esta criança* out of the complex NP in (18), the result turns out to be ungrammatical. Consider

(19) (a) *É esta criança que não há ninguém que queira adotar.

'It is this child that there is nobody who is willing to adopt'.

(b) *Esta é a criança que não há ninguém que queira adotar.

'This is the child that there is nobody who is willing to adopt'.

(c) *Quem não há ninguém que queira adotar?

'Who there is nobody who is willing to adopt?'

Both (5) and (8) predict, correctly, the ungrammaticality of (19). Portuguese does not care if the head NP of a relative clause is 'semantically transparent' or not. Once there is a relative clause in Portuguese, nothing can be moved out of it. In the strongest interpretation of (9), the sentences of (19) can be taken as examples of violations to the *Thematic Constraint on Relative Clauses*.

Let us consider now the second aspect of the complex NP, namely those cases of the *fact-S* class. Consider

(20) João deu a sugestão de que Maria devia tomar leite.

'João made the suggestion that Maria should drink milk'.

Both (5) and (8) predict that we cannot move the NP *leite* out of the complex NP. But the sentences of (21) are grammatical:

(21) (a) Foi leite que João deu a sugestão de que Maria devia tomar.

'It was milk that João made the suggestion that Maria should drink'.

- (b) Este é o leite que João deu a sugestão de que Maria devia tomar.
 'This is the milk that João made the suggestion that Maria should drink.
- (c) O que João deu a sugestão de que Maria devia beber?
 What did João give the suggestion that Maria should drink?

The sentences of (21) contradict (5) and (8). Concerning (9), it makes no predictions whatsoever about (21). Cases like (21) can be multiplied. Consider (22) and (23).

- (22) (a) Bonnie mencionou o fato de que Clyde queria roubar este asilo.
 Bonnie mentioned the fact that Clyde wanted to rob this asylum.
- (b) Este é o asilo que Bonnie mencionou o fato de que Clyde queria roubar.
 This is the asylum that Bonnie mentioned the fact that Clyde wanted to rob.
- (c) Foi este asilo que Bonnie mencionou o fato de que Clyde queria roubar.
 It was this asylum that Bonnie mentioned the fact that Clyde wanted to rob.
- (d) Que asilo Bonnie mencionou o fato de que Clyde queria roubar?
 Which asylum did Bonnie mention the fact that Clyde wanted to rob?
- (23) (a) João tem medo que Maria compre este elefante.
 João 'has fear' (is afraid) that Maria buys this elephant.
- (b) É este elefante que João tem medo que Maria compre.
 It is this elephant that João 'has fear' (is afraid) that Maria buys.
- (c) Este é o elefante que João tem medo que Maria compre.
 This is the elephant that João 'has fear' (is afraid) that Maria buys.
- (d) O que João tem medo que Maria compre?
 What is João afraid of that Maria buys?

In the sentences (24) and (25), the underlined NP's can be equally extracted out of the complex NP:

(24) João tem tendência a caçar *ratos*.
João has a tendency to chase *mice*.

(25) João dava a impressão de querer *dinheiro*.
João gave the impression that he wanted *money*.

Our task here will be to try an explanation for these facts.

2. A Tentative Explanation

As we have seen, neither (5) nor (8) can account for the facts from (20) to (25) in Portuguese. On the other hand, (9) is also useless since it is a constraint on relative clauses and not on the *fact-S* class of sentences. We are left with two possible ways to try an explanation for these facts. First, we can think about motivating two different syntactic structures, one for relative clauses and another one for the *fact-S* class. I do not know any good reason to propose two different syntactic structures here. In fact, what I know of is a proposal (cf. Stahlke, 1976) by which the underlying structure of some relative clauses, those with *that*, are very similar to the structure underlying *fact-S*, to the extent that this *that* is considered a complementizer. Second, we could think of a non-syntactic difference to explain these facts. We will pursue the second alternative since we cannot think of any reasonable reason to differentiate the two cases syntactically.

Before attacking this problem directly, let us consider a different aspect of the complex NP. Suppose a sentence like

(26) Maria comprou um carro vermelho.
Maria bought a car red (a red car).

where *um carro vermelho* is an NP which dominates Det + N + Adj. Any transformation applied to the NP *um carro vermelho* will be applied to the whole string Det + N + Adj. We cannot, e.g., apply a transformation to *carro* and leave *um* and *vermelho* unaffected. Consider the sentences in (27),

- (27) (a) *O que Maria comprou (um) vermelho?
What did Maria buy (a) red?
(b) *Um carro, Maria comprou vermelho.
A car, Maria bought red.
(c) *Foi um carro que Maria comprou vermelho.
It was a car that Maria bought red.

where the adjective *vermelho* was unaffected by the transformations. Now, if also the adjective *vermelho* had been affected, together with *um carro*, we would get the grammatical sentences in (28) :

- (28) (a) O que Maria comprou?
What did Maria buy?
- (b) Um carro vermelho, Maria comprou.
A red car, Maria bought.
- (c) Foi um carro vermelho que Maria comprou.
It was a red car that Maria bought.

Consider now sentence (29),

- (28) João mandou flores a Maria.
João sent flowers to Maria.

where the VP dominates V — NP — PP. Unlike (26), a transformation can be applied to the VP in (29), affecting just one element. For instance, we can question *flores* alone, or *a Maria* alone, as in (30) and (31), respectively:

- (30) O que João mandou a Maria?
What did João send to Maria?
- (31) A quem João mandou flores?
To whom did João send flowers?

Also, Topicalization or Pseudo-Cleft could be applied either to *flores* or to *a Maria*, without any obligation of being applied to both at the same time. But, why this is so? In (26) the N *carro* is the theme, in Kuno's sense, of the NP *um carro vermelho* (and *um* and *vermelho* are 'saying something about' *carro*), whereas in (29) *flores* and *a Maria* are not in a theme relation to each other, i.e., one is not the theme of their sum and also no one of them, without intonational recourse, is the theme of the VP. We will return to this point later. Consider now (32) and (20):

- (32) João leu o livro que estava sobre a mesa.
João read the book that was on the table.
- (20) João deu a sugestão de que Maria devia tomar leite.
João made the suggestion that Maria should drink milk.

In (32) the NP *o livro* is the theme of the relative clause. Like in (26), in (32) the theme cannot be separated from the other constituents which are, together with the theme, under the same category. Consider the sentences of (33):

- (33) (a) *O que João leu estava sobre a mesa?
What did João read that was on the table?
- (b) *O livro João leu que estava sobre a mesa .
The book João read that was on the table.
- (c) *Foi um livro que João leu que estava sobre a mesa
It was the book that João read that was on the table.

But if we apply the same transformations to the relative clause in (32) plus its head noun, the result is grammatical.

- (34) (a) O que João leu?
What did João read?
(b) O livro que estava sobre a mesa João leu.
The book that was on the table João read.
(c) Foi o livro que estava sobre a mesa que João leu.
It was the book that was on the table that João read.

Notice that Portuguese, unlike English, does not allow relative clauses extraposed from the complex NP. Consider the sentences of (35), which are the equivalents to (6a) and (7a).

- (35) (a) O homem que comprou o carro pulou a cerca.
The man who bought the car jumped over the fence.
(b) *O homem pulou a cerca que comprou o carro.
The man jumped over the fence who bought the car.

(35b) has an anomalous reading, where it is said that *the fence bought the car*. But if we find a parallel between (26) and (32), concerning the impossibility to extract certain elements alone, the same parallel is not found between (29) and (20). In (20) the NP *a sugestão* and the complement *que Maria devia tomar leite* are not in a theme relation to each other, i.e., the complement is not a statement about *a sugestão*. It is simply its syntactic complement (as it would be the syntactic complement of the verb *sugerir* 'to suggest' and not a statement about this verb). Even though, it is not possible to separate *a sugestão* from *que Maria devia tomar leite*. Consider

- (36) (a) *A sugestão João deu de que Maria devia tomar leite.
The suggestion João made that Maria should drink milk.
(b) *Foi esta sugestão que João deu de que Maria devia tomar leite.
It was this suggestion that João made that Maria should drink milk.

On the other hand, Topicalization and Pseudo-Cleft can be applied to the whole complex NP. Consider

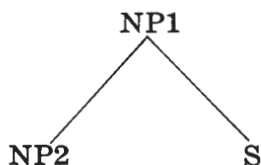
- (37) (a) A sugestão de que Maria devia tomar leite João deu.
The suggestion that Maria should drink milk João made.
(b) Foi a sugestão de que Maria devia tomar leite que João deu.
It was the suggestion that Maria should drink milk that João made.

And this is another difference between Portuguese and English. Consider the following sentence in Ross (1968), repeated here as

- (38) A proof was given that the claim that John had lied had been made.

where *a proof* is separated from its complement *that the claim...* by an intervening *was given*. The Portuguese correspondent to (38) would have to have *was given* in last position. So, up to now we know two things about the complex NP in Portuguese. First, whenever we have the configuration of (1)

(1)



no transformation can separate *NP2* from *S*. Second, whenever (1) is related to relative clauses, no element contained in *S* can be moved out of *NP1*, but if (1) is related to the *fact-S* class of sentences, then an element contained in *S* can be extracted out of *NP1*. The first aspect of the complex NP can be accounted for by the following constraint:

- (39) No transformation can separate a *NP1* from a *S1* if both of them are directly dominated by another *NP2*.

(39) is being proposed here as a language-specific constraint for Portuguese. It blocks explicitly a possible extraposition of *S* out of *NP1* in (1), and also any movement of *NP2* out of *NP1*. There is no need to mention that the *NP1* in (39) has to be a lexical item since Portuguese has nothing corresponding to the dummy *it* of English. Sentences like (33), (35b) are ruled out by (39).

For the second aspect of the complex NP we will propose a tentative constraint in terms of *theme*. As we have seen (Cf. (32) and (20)), heads of relative clauses are their theme, whereas lexical heads of *fact-S* sentences are not the theme of their complements. Suppose we have the following constraint:

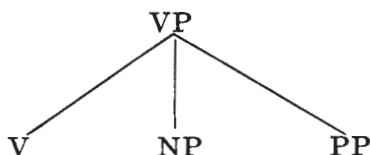
- (40) If in a structure ... [...B...C...], B is theme of C,

A

then no element contained in C can be moved transformationally out of A.

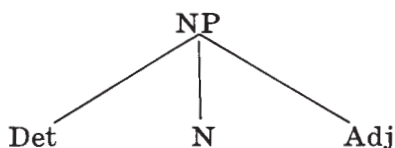
(40) predicts that sentences (15), (16), (17) and (19) are ungrammatical since their theme was crossed over by an element contained in the relative clause. On the other hand, (40) does not block (21) - (23). And that is the result we want. Now, we could like that our use of the notion of theme had a broader scope. It would sound a little ad hoc if *theme*, as a constraint, worked only for the cases we just mentioned. So, consider the facts we pointed in (26) - (28), on the one hand, and (29) - (30) on the other hand. We pointed out that in the NP *um carro vermelho* 'a red car' the adjective *vermelho* 'red' is stating something about the car. In this sense, *carro* 'car' is the theme of the whole NP. We said, too, that in the VP *mandou flores a Maria* 'sent flowers to Maria' we cannot say that the PP *a Maria* is stating something about *flores*, or vice versa, and so no one of them is the theme in the VP (either one could be the theme of the whole S, either by receiving a stronger stress or by being moved to the first position, for instance, in the sentence). Why is it possible to extract either the NP or the PP out of the VP in a structure like

(41)



while it is not possible to extract the N out of the NP in a structure like (42)?

(42)



We propose (43) as a new version of (40) to account for these facts too:

- (43) If in a structure ... [...B...C...] B is the theme of C, then
- A
- a) no element contained in C can be moved transformationally out of A,
 - b) nor B can be moved out of A leaving C behind it.

Given (43), the sentences of (27) can be blocked. Also, sentences (33) and (35b) are doubly blocked: by (39) and by (43b).

Let us consider now another point. Suppose that the transformational modifications of syntactic structures are not simply another way to say the same thing, but rather that these modifications convey another type of information. Let us take as example (20) and its passive in (44) :

(20) João deu a sugestão de que Maria devia tomar leite.
João made the suggestion that Maria should drink milk.

(44) A sugestão de que Maria devia tomar leite foi dada por João.

The suggestion that Maria should drink milk was given by João.

Let us say that the Passive is a syntactic device to assign a constituent a thematic character (notice that while there are agentless passives, there are no passives without a surface subject). Now, if we extract *leite* 'milk' out of the complex NP in (44), as we did in (21), the result is

(45) (a) *Foi leite que a sugestão de que Maria devia tomar foi dada por João.

It was milk that the suggestion that Maria should drink was given by João.

(b) *Este é o leite que a sugestão de que Maria devia tomar foi dada por João.

This is the milk that the suggestion that Maria should drink was given by João.

(c) *O que a sugestão de que Maria devia tomar foi dada por João?

What the suggestion that Maria should drink was given by João?

and (43) is unable to block these sentences. In (44) the whole sentence satisfies A in (43), the complex NP is B and C is *foi dada por João*, B is the theme, but in (45) no element contained in C can be moved out of A ((43a) is not being violated), nor can B be moved out of A ((43b) is not being violated). It seems that elements in the theme cannot be moved out of the theme. Our theme constraint will be restated as

(46) *Theme Constraint on Movement Rules*

If in a structure ... [...B...C...] N is the theme of C, then A

a) no element contained in C can be moved out of A,

b) nor B can be moved out of A leaving C behind it,

c) nor any element contained in B can be moved out of B.

3. *The Scope of the Theme Constraint*

In this final section we will see if our theme constraint can be used to account for other facts. Certainly (46) (and (39)) explains more facts involving complex NP's in Portuguese than either (5), (8) or (9). But there are other constraints raised in Ross (1968) which provide the best test for (46). One of them is the *Sentential Subject Constraint*. Ross points out that noun phrases in (47a) and (47c) can be relativized, but not in (47b), as shown in (48):

- (47) (a) The reporters expected that the principal would fire some teacher.
(b) That the principal would fire some teacher was expected by the reporters.
(c) It was expected by the reporters that the principal would fire some teacher.
- (48) (a) The teacher who the reporters expected that the principal would fire is a crusty old battle-ax.
(b) *The teacher who that the principal would fire was expected by the reporters is a crusty old battle-ax.
(c) The teacher who it was expected by the reporters that the principal would fire is a crusty old battle-ax.

To account for this Ross proposes the *Sentential Subject Constraint*, restated here as

- (49) No element dominated by an S may be moved out of that S if that S node is dominated by an NP which is itself immediately dominated by S.

As we have shown before (cf. (44), (45)), this can be accounted for by (46). In (47b) the *that*-clause is the theme. In (48b) the NP *the teacher* is being moved out of the theme, violating (46c).

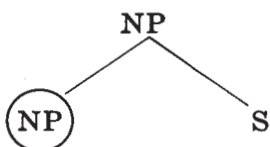
Another Island Constraint proposed by Ross is the *Left Branch Condition*, restated here as

- (50) No NP which is the leftmost constituent of a larger NP can be reordered out of this NP by a transformational rule.

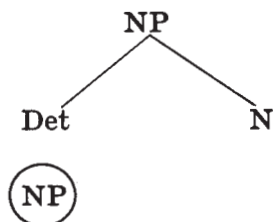
(50) was proposed to account for two sets of facts. First, the impossibility to question or relativize the head NP of a relative clause and, second, the impossibility to question or relativize NP's, which are dominated by Det, out of the NP which dominates Det.

So, the circled NP in (51) and (51') cannot be extracted out of the underlined NP:

(51)

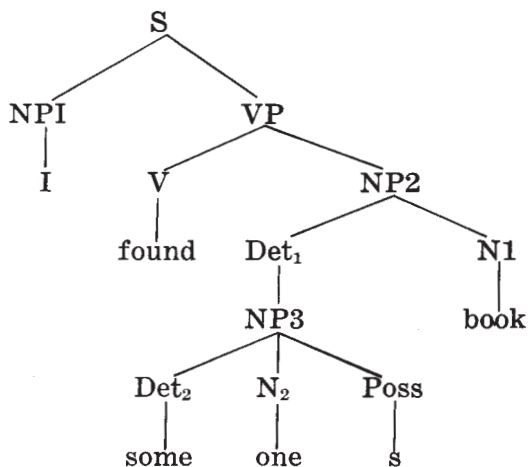


(51')



As we have seen before (cf. (33)), (46b) blocks the extraction in (51). Now (51') is a case involving genitives. Ross' example is

(52)



and (50) blocks (53),

(53) *Whose did you find book?

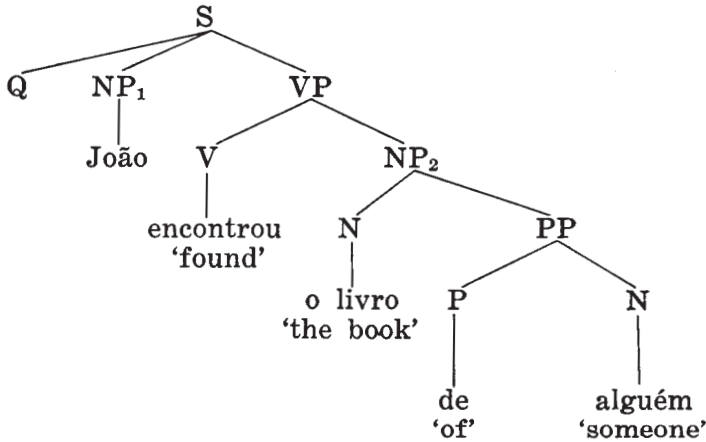
where NP₃ of (52) was moved out of NP₂, and allows

(54) Whose book did you find?

where NP₂ is being questioned. Our (46a) predicts exactly the same facts. The N *book* in (52) is the theme of NP₂, in Kuno's sense. On the other hand, (50) cannot account for similar facts of Portuguese, where in a case like (52) Det₁ and N₁ have their positions

reversed under NP₂ (in fact there will be a PP instead of Det₁), as in

(55)



Since N's under PP can be questioned, there is nothing to prevent us from deriving

(56) *De quem João encontrou o livro?

Whose did João find the book?

Again, (46) blocks (56). What we can have is a question on NP₂ of (55), deriving the sentences of

(56') O livro de quem João encontrou?

Whose book did João find?

Finally, we would like to refer to Emond's work (1969). One of the properties of the Root-Transformations is that just one can be applied in the same root sentence. The reason why the sentences of (57) are ungrammatical is that more than one Rot-Transformation was applied in the root sentence:

(57) (a) *John, away he ran. (Directional Adverb Preposing + Left dislocation)

(b) *Never these steps did I sweep with a broom.

(Topicalization + Negated Constituent Preposing)

just to give two examples. A complete list, from which (57) was extracted, can be found in Emonds (1972). Hooper and Thompson (1973) suggest that the Root-Transformations insert emphasis in the moved constituent. Consider

(58) (a) Away ran John.

(b) These steps I never swept with a broom.

where *away* and *these steps* are the peak of prominence in their respective sentences. But is there any reason why (57) should be ungrammatical? To say that two (or more) Rot-Transformations cannot be applied in the same root sentence is a statement derived from observation, but it does not explain why this is so. In (58), *away* and *these steps* are the theme (in the sense we are giving to the term) of *a* and *b*, respectively. Take, for example, (58b), where *These steps I never swept with a broom* satisfies A in (46), and where *these steps* satisfies B in (46), and where *I never swept with a broom* satisfies C in (46). This being the case, (57b) (as well as (57a)) is ungrammatical because (46a) is being violated, since an element of C, *never*, is being moved out of A.

4. Conclusion

In this paper we tried to provide a constraint on movement rules in terms of *theme*, namely (46), which works better for Portuguese than either (5) or (8). Also, it covers more facts than (9). (46) is not being proposed as a universal constraint. We do not have arguments to propose this. On the other hand we have seen that (46) has a larger scope, and goes beyond those cases involving the notion of the complex NP. Also, (46) can be used to provide an explanation for the impossibility of applying more than one Root-Transformation in the same root sentence. The constraint in (46) is a tentative one and certainly will have to be further refined.

References

- EMONDS, J. 1969. "Root and Structure Preserving Transformations." Ph.D. dissertation, MIT, Cambridge, Mass.
- EMONDS, J. 1972. "A reformulation of certain syntactic transformations", in S. Peters (ed.) *Goals of Linguistic Theory*, pp. 21-62. Englewood Cliffs: Prentice Hall.
- HALLIDAY, J. B., and S. A. Thompson. 1973. "On the applicability of Root-Transformations", *Linguistic Inquiry* 4. (4) :465-497.
- KOUTSOUDAS, A. 1973. "Extrinsic order and the Complex NP Constraint", *Linguistic Inquiry* 4. (1) :69-81.
- KUNO, S. 1976. "Subject, theme, and the speaker's empathy", in C. N. Li (ed.), *Subject and Topic*, 419-444. Englewood Cliffs: Prentice Hall.
- ROSS, J. R. 1968. *Constraints on Variables in Syntax*, mimeograph, Indiana University Linguistics Club, Bloomington, Indiana.
- STAHLKE, H. F. W. 1976. "Which that", *Language* 52. (3).