

YORUBA RELATIVISATION AND THE CONTINUOUS SEGMENT PRINCIPLE*

Nike S. Lawal
University of Ilorin

This paper examines the strategies for relativisation and the Noun Phrase accessibility hierarchy and constraints in Yoruba [Keenan and Comrie 1977]. The various positions relativisable are examined. It was found that contrary to what Keenan and Comrie thought, Yoruba relativises all positions except the Object of Comparative. Attention is also focussed on the status of the coreferential pronoun found in subject relativisation. From the presence of this pronoun it appears as if Yoruba violates the continuous segment principle. However, our analysis shows that the pronoun is a surface structure phenomenon which obscures the underlying strategy for relativisation. The Yoruba data therefore supports the Hierarchy Constraints.

1. Introduction

One important issue which often comes up in the description of relative clauses (RC) in various languages is the relativisation strategy proposed by Keenan and Comrie [1977]. Keenan and Comrie observed that languages vary with respect to the way relative clauses (RC) are formed. They also noted that even within a single language there is often more than one distinct way of forming a RC. The different ways of forming RC is what they refer to as different relative clause forming Strategies. Different strategies differ with regard to which NP positions they can relativize. Their data covers fifty languages. Essentially, their observation distinguishes two main parameters of variation. The first concerns the position of RCs with respect to the main

*The author wishes to acknowledge the comments and helpful criticisms of the editor of SAL on an earlier version of this paper. This work was supported by a University of Ilorin Senate Research Grant.

clause. For this, three distinct types are recognised. They are postnominal, prenominal, and internal. A postnominal RC is one which is positioned immediately following the matrix constituent within the main clause, a prenominal RC is positioned immediately to the left of its head, while an internal RC is embedded internally with the main clause, and it has no matrix constituent.

Their second parameter classifies RC according to whether they are case-coding or non-case coding. A RC is case-coding if it overtly marks the grammatical role of the relative within the RC. This may be achieved either by morphological marking, the use of pronouns, prepositions or postpositions, or by position. They also observed that some syntactic positions are more accessible to relativisation than others. This they called the accessibility hierarchy. First in the hierarchy is subject position, followed by direct objects, indirect objects, obliques, and lastly, objects of comparison.

Keenan and Comrie proposed a number of universal principles which make crucial reference to this hierarchy. The most important of these as far as Yoruba is concerned is what is described as "the continuous segment principle", and I quote:

"For unmarked simple sentences, any relative clause-forming strategy must operate on a continuous segment of the accessibility hierarchy."

This in effect means that a given RC-forming strategy can be used to relativize only continuous segments of the hierarchy. The continuous segment principle has been tested on many languages, including Yoruba. This paper discusses the proposal as it relates to Yoruba data. The focus is on the coreferential subject pronoun in Yoruba subject relatives. But first we consider the range of constituents relativizable in the language.

2. Range of Constituents Relativizable in Yoruba

Yoruba was one of the fifty languages examined by Keenan and Comrie. According to Keenan and Comrie, Yoruba relativises only subject (SU), Direct Object (DO), and Genitive (Gen) positions; the Indirect Object (IO), Oblique (Obl.), and Object of Comparison (Ocomp) are not relativizable. These three positions are therefore marked with asterisks in their paper. However, the

following sentences indicate that Yoruba relativizes all positions:

- (1) ọ̀kùnrin tí ó ra aṣọ̀ tí dé
man that he bought cloth has come
'the man who bought the cloth has come'
- (2) aṣọ̀ tí ọ̀kùnrin náà rà dára
cloth that man the bought good
'the cloth that the man bought is nice'
- (3) ọ̀mọ̀ tí ọ̀kùnrin náà ra aṣọ̀ fún tí dé
child that man the bought cloth for has come
'the child that the man bought the cloth for has come'
- (4) àdà tí mo fi gé igi mú
cutlass that I with cut tree sharp
'the cutlass that I cut the tree with is sharp'
- (5) iná tí mo gbé ọ̀bẹ̀ kà tí kú
fire that I put soup on has died
'the fire on which I put the soup is dead'
- (6) obìnrin tí awọn olè gbé mọ̀to rẹ̀ tí kú
woman that plural thief took car her has died
'the woman that thieves stole her car has died'
- (7) obìnrin tí ọ̀kùnrin náà ba lọ̀ tí de
woman that man the with go has arrived
'the woman that the man went with has arrived'
- (8) ọ̀mọ̀ tí ọ̀kùnrin náà ga jù̀ tí dé
child that man the tall exceed has arrived
'the child that the man is taller than has arrived'

The examples (1) to (8) illustrate the full range of constituent types that can be relativised: (1) is an example of a subject being relativised; (2) is that of direct object; (3) is an indirect object; (4), (5), and (7) are obliques; (6) is a genitive; (8) is object of comparison. The three positions which have been a subject of dispute are the indirect object, oblique, and object of comparison. Keenan and Comrie for instance do not regard the Yoruba RC like (3), (4), (5), (7), and (8) as instances of IO, Oblique, and OComp con-

structions respectively. Therefore these positions are marked with asterisks in their table of relativisable positions. Their reason is that the words *fún*, *bá*, *ka*, and *jù* which precede the relativised NPs are verbs in serial verbal constructions and not prepositions. If their interpretation is correct the implication is that the NPs which are complements to these words would be direct objects. Thus there would be no special IO, Oblique, and OComp constructions in the language. We consider Keenan and Comrie's claim to be valid only in the case of the OComp position (8). The word *jù* in (8) is in all respects a verb. However, this is not the case with the words *fún*, *fi*, *ka*, and *bá* in (3), (4), (5), and (7) respectively.

It is true as some linguists have said, that these words were historically verbs, and it is this which probably influenced Keenan and Comrie in their analysis. However, the concern here is not and should not be with the diachronic status of these words. What is at issue here is the present status of these words, and linguists are now reanalysing these words as prepositions or case markers. An explicit account of the present status of these words can be found in Awobuluyi [1978] and Lawal [1986] amongst others and will not be discussed here.

3. Pronoun Retention and the Continuous Segment Principle

According to the Continuous Segment Principle quoted in section 1 above, a given RC-forming strategy can be used to relativise only continuous segments of the accessibility hierarchy. For example, if a language has two relativisation strategies, a case coding strategy and a non-case coding strategy, and if the case coding strategy is used to relativise only subjects, indirect objects, and Obliques, this would in effect leave a gap, i.e. the direct object position relativised by a different strategy would destroy the continuity.

Case is coded either by means of a personal pronoun or a stranded preposition within the relative clause. Now consider the Yoruba RCs below:

- (9) *bàbá* [tí o ra bata]
 father that he bought shoes
 'father who bought shoes'

- (10) bàtà [ti [baba ra Ø]
shoe that father bought
'the shoe that father bought'
- (11) ọmọ [ti [bàbá ra bàtà fún Ø]]
child that father the bought shoe for
'the child that father bought shoes for'
- (12) àdá [ti [mo fi Ø be isu]]
cutlass that I with cut yam
'the cutlass that I cut the yams with'
- (13) iná [ti [mo gbé ọbẹ ka Ø]]
fire that I put soup on
'that fire that I put the soup on'
- (14) ọkùnrin [tí [obìnrin náà ba Ø lo]]
man that woman the with went
'the man that the women went with'
- (15) ọkùnrin [ti [awon ole ji mótò rè]]
man that they thieves stole car his
'the man whose car the thieves stole'
- (16) ọkùnrin [ti [mo ga ju Ø]]
man that I tall surpass
'the man that I am taller than'

Starting with (9) we can see there is no missing subject in the RC. The relativized constituent which is subject is represented in the RC in the form of a personal pronoun. In (10) on the other hand the relativized NP which is a direct object is not represented in the RC; the NPrel has been deleted. The same applies to all the other positions except the genitive position (15) where the NPrel is pronominalised. We have decided to leave out example (16) in this analysis because we agree with Keenan and Comrie that it is not a true object of comparison. The NPrel here has the status of a direct object.

Now the indication from this analysis is that Yoruba has two strategies, a case coding strategy (pronominalisation) for subjects and genitives and a non case coding strategy (deletion) for direct objects, Indirect Objects, and Obliques. In the latter, there is no nominal element in the RC that clearly ex-

presses which NP position is being relativized.

This strategy appears to constitute a counterexample to the Hierarchy Constraint. When we subject it to the Continuous Segment Principle the RC forming strategies in Yoruba do not operate on a continuous segment. It shows a gap along the segment of the hierarchy brought on by the use of a different strategy for direct objects, indirect objects, and obliques.

The Yoruba data can also be interpreted in another way, i.e. depending on how we interpret the notion "case coding". If we take case coding to include stranded prepositions as some linguists do, then Yoruba RC strategy can be said to be case coding for all positions except the direct object. The stranded prepositions in this case are *fún*, *fi*, *kà*, and *bá*. Whichever interpretation we choose, the implications appear to be the same, which is that there is a gap along the segment. With this second interpretation the gap would be brought on by the use of a different strategy for direct objects which is second on the accessibility hierarchy.

We shall in this paper adopt the first approach, which means case coding for subjects and genitives and non-case coding for all other positions viz, direct objects, indirect objects, and obliques.

This violation of the Continuous Segment Principle in Yoruba was noticed earlier by Keenan and Comrie themselves. To explain the "violation" Keenan and Comrie [1977] claimed that the pronoun found in Yoruba subject RC is not a pronoun at all but a case of agreement. This in effect would mean that subject, direct object, indirect object, and oblique positions are all non-case coding. Only the genitive which is lowest on the Yoruba hierarchy would then be case coding, thus preserving the continuity.

4. The Case Against the Pronoun *ó* as an Agreement Marker

We disagree with the claim that the pronoun in subject relatives is an agreement marker. Firstly, there is no evidence that supports such a claim. For instance, Hausa, with which Keenan and Comrie by implication compare Yoruba, is very much unlike Yoruba. Yoruba for example does not retain a full pronoun in simplex sentences with a subject NP as Hausa does. Thus the following are unacceptable in Yoruba:

- (17) a. *Olú ó wá
Olu he came
- (18) a. *bàbá ó lọ
father he went

The correct sentences are as follows:

- (17) b. Olú wá
Olu came
- (18) b. bàbá lọ
father went

Compare the above with an equivalent sentence in Hausa:

- (19) a. Musa ya zo
Musa he came
- b. *Musa zo
Musa came

Secondly, there is no evidence of full subject NPs accompanied by Clitic pronouns in Yoruba as justified in Givón [1976]. The pronoun present in Yoruba subject RC is a full pronoun, not a Clitic. This is illustrated with the sentence below:

- (20) ó ra bàtà 'he/she bought shoes'
he/she bought shoes
- (21) ó ñ sùn 'he/she is sleeping'
he/she prog sleep

The analysis used for Hausa and ergative languages like Tongan does not work for Yoruba.

Another linguist who also disagrees with Keenan and Comrie is Stalhke [1976]. Stalhke is of the opinion that the element ó found in Yoruba subject RC is not an agreement morpheme as claimed by Keenan and Comrie. Stalhke argues that this ó occurs with both plural and singular subjects in focus constructions. His examples are reproduced as (22a-d) below:

- (22) a. èmi ni ó lọ 'it is I that went'
I emph he/she went
- b. ìwọ ni ó lọ 'it is you that went'
you emph he/she went

- c. àwa ni ó lọ 'it is we that went'
we emph he/she went
- d. àwọn ni ó lọ 'it is they that went'
they emph he/she went

Stalhke argues that if ó were an agreement morpheme, this would be difficult to explain. As he says, "It would mean claiming that the third person singular pronoun can agree with any other person/number combination." Such a claim, he points out, cannot be supported from any other area of Yoruba syntax, for example, plural examples are ungrammatical with just number agreement. Consider the examples below from Stalhke:

- (22) e. *èyin ni ó lọ
you (P1) emph they went
- f. *àwa ni wón lọ
we emph they went

The above are ungrammatical because there is no person agreement. The pronouns must agree both in number and person. Stalhke's conclusion therefore is that ó is not an agreement morpheme.

Although Stalhke's argument is against ó as an agreement marker, it also seems to be an argument against ó as a subject pronoun. Stalhke states, "It would be an odd subject pronoun which remained constant across this set of sentences."

According to Stalhke, ó is neither a subject pronoun nor an agreement morpheme. But what is it? Unfortunately, Stalhke does not seem to have an answer. That ó is a full subject pronoun is not in doubt as it occurs in simple sentences like ó ra bàtà fún mi 'he bought shoes for me'. It is its status in RC that is in doubt.

Apart from RC there is only one other construction in which a pronoun follows a full subject NP. This is focus construction. This has been used as evidence in favour of analysing ó as an agreement morpheme. However, it should not be surprising that the ó occurs also in focus constructions. Studies have shown that many languages employ the same strategies for both focus formation and RC formation [Schachter 1973, Madugu 1982]. In all the languages ex-

amined by Schachter, the rules for forming RC and focus clauses are alike with regard to such details as the substitution of an appropriate personal pronoun for one of the NPs in the sentence or the insertion of an invariant marker at the beginning of the clause and an optional deletion of the relative pronoun in certain cases.

Yoruba is no exception in this regard. Thus, to form focus clauses in Yoruba the same rules as used for relative clauses are employed, except that the markers inserted at the beginning of the two clauses are different: the focus marker is *ni* while the relative marker is *ti*. The following examples illustrate this.

- (23) a. òrẹ́ mí tí ó ra bàtà (Subject RC)
friend my that he bought shoes 'my friend that bought shoes'
b. òrẹ́ mí ní ó ra bàtà (Subject focus)
friend my emph he bought shoes 'it is my friend that bought shoes'
- (24) a. òrẹ́ mí tí mò rí (Object RC)
friend my that I saw 'my friend that I saw'
b. òrẹ́ mí ní mò rí (Object focus)
friend my emph I saw 'it is my friend that I saw'
- (25) a. òrẹ́ mí tí mò ra bàtà rẹ́ (Genitive RC)
friend my that I bought shoes her 'my friend whose shoes I bought'
b. òrẹ́ mí ní mò ra bàtà rẹ́ (Genitive focus)
friend my emph I bought shoes her 'it is my friend whose shoes I bought'
- (26) a. òrẹ́ mí tí mò ga jù (OComp RC)
friend my that I tall pass 'my friend that I am taller than'
b. òrẹ́ mí ní mò ga jù (OComp focus)
friend my emph I tall pass 'it is my friend I am taller than'

Just as in the RC, a pronoun is retained when Subject NP and Genitive NP are being focussed, while all other positions do not retain pronouns. These similarities between focus sentences and RC correlate with their semantics: both constructions are said to have a common semantic property which is "foregrounding of one part of a sentence at the expense of the rest" [Schachter 1973]. Evidence based on a comparison with focus sentences in the language does not therefore constitute independent evidence for the pronoun as an agree-

ment marker. It would be odd if the pronoun were to be absent in focus constructions since the same strategy is employed for both constructions. Before we conclude this section we want to draw attention to another type of subject relativization which constitutes additional evidence against the pronoun *ó* as a case marker. Contrast the (a) sentences below with the (b) sentences:

- (27) a. *ìwọ obìnrin tí ó gbọ gèḗsì*
 you woman that she hears English
 'you woman who understands English'
- b. *ìwọ obìnrin tí ọ gbọ gèḗsì*
 you woman that you hear English
 'you woman who understands English'
- (28) a. *àwa obìnrin tí ó gbọ gèḗsì*
 we women that she hear English
 'we women who understand English'
- b. *àwa obìnrin tí a gbọ gèḗsì*
 we women that we hear English
 'we women who understand English'
- (29) a. *àwọn obìnrin tí ó gbọ gèḗsì*
 them women that she hear English
 'those women who understand English'
- b. *àwọn obìnrin tí wọn gbọ gèḗsì*
 them women that they hear English
 'those women who understand English'
- (30) a. *èyin obìnrin tí ó gbọ gèḗsì*
 you (pl) women that she hear English
 'you (pl) women who understand English'
- b. *èyin obìnrin tí ẹ gbọ gèḗsì*
 you (pl) women that you (pl) hear English
 'you (pl) women who understand English'

In the (b) sentences we have the full range of pronouns, i.e. we have 1st, 2nd, and 3rd persons as heads of the RC, and the pronouns agree in number and person with their antecedents. This is unlike the (a) sentences where the 3rd

person singular pronoun *ó* is used for all persons and number. The fact that the surface subject in the (a) sentences of (27-30) does not need to show agreement even strengthens the claim that the pronoun *ó* is not an agreement or case marker.

4. The Status of the Pronoun in Subject Relatives

To account for the pronoun *ó* in subject relatives it is necessary to examine the facts of the language. Perlmutter [1970] observed that for various reasons a class of surface structure constraints or input conditions must be available as one of the devices used by grammars to delimit the class of grammatical sentences, in other words, all languages have certain well-formedness conditions which determine which sentences are well formed and which are not. English for instance, is said to have such a constraint, which can be stated, "No non-imperative sentence which does not have a subject in its surface structure is well formed." This is a surface structure constraint about English.

Yoruba like other languages has surface structure constraints. One such constraint is that "no top-most sentence in Yoruba may be verbless." Another is "no tensed sentence or clause may be subjectless" [Lawal 1985]. We must note, however, that this constraint does not apply to negative sentences where the subject is a 3rd person pronoun. In Yoruba when the 3rd person subject pronoun of a sentence comes in contact with the negator *kò*, the pronoun gets elided. However, the sentence would still be understood as containing a subject, even though the subject no longer has an independent existence. The subject pronoun more or less fuses with the negator *kò*. Such sentences cannot in fact be said to be subjectless. The following is an example of this construction:

(31) *kò* *pá á* 'he/she didn't kill it'
 he/she not kill it

The negator *kò* incorporates the subject. Thus the sentence cannot be interpreted as being subjectless. This phenomenon is restricted to the negator *kò* only. Thus (32) is ungrammatical:

- (32) *ò pa á
not kill it

We can now account for the pronoun in subject RC. The pronoun is there to satisfy the above Yoruba surface structure constraint and has nothing to do with the underlying strategy for relativization. It simply obscures the underlying strategy of deletion of the coreferential NP, the same strategy as used for direct objects.

5. Conclusions

We have shown firstly, that Yoruba relativizes all NP positions and secondly, that Yoruba obeys the Continuous Segment Principle. The pronoun in subject RC is a surface structure phenomenon which obscures the underlying strategy for relativization.

REFERENCES

- Awobuluyi, A.O. 1978. *Essentials of Yoruba Grammar*. Ibadan: Oxford University Press, Nigeria.
- Givón, T. 1976. "Topic, pronoun, and agreement." In C. Li (ed.), *Subject and Topic*, pp. 149-188. New York: Academic Press.
- Keenan, L. and B. Comrie. 1977. "Noun phrase accessibility and universal grammar." *Linguistic Inquiry* 8:63-99.
- Lawal, S.N. 1985. "Why verbs do not gap in Yoruba." *Journal of African Languages and Linguistics* 7:155-162.
- Lawal, S.N. 1986. "The classification of the Yoruba serial verbal construction." Submitted to *JOLAN*.
- Madugu, I.S. 1982. "The na-na construction in Nupe." *JOLAN* 1:34-45.
- Perlmutter, D. 1970. *Deep and Surface Structure Constraints in Syntax*. New York: Holt, Rinehart and Winston.
- Schachter, P. 1973. "Focus and relativization." *Language* 49:19-49.
- Stalhke, H. 1974. "Pronouns and islands in Yoruba." *Studies in African Linguistics* 5:171-202.

