

EMPTY OPERATOR RAISING IN KITHARAKA*

Carolyn Harford
University of Zimbabwe

The Kenyan Bantu language Kitharaka (E.54) has a biclausal construction in which a verb with an athematic subject takes a finite complement. A referential NP in the matrix subject position may acquire its thematic role through linkage with an argument at a lower level of embedding. The paper presents evidence that this construction, referred to as Empty Operator Raising, involves empty operator movement, although it is also semantically similar to Raising-to-Subject. If the matrix subject is distinct from the subject of the nearest embedded clause, this clause must contain a clefted NP. The author proposes that this clefting is a distancing strategy to prevent an empty operator in SPEC/CP from binding the embedded subject position, to avoid forming an illicit chain with the matrix subject.

* I would like to thank the following speakers of Kitharaka for their willing and patient assistance: Kaburu Kaimba (Gatunga, Mũkothima), Mũtugi Kibaara (Mũkothima), Nyaga Mzalendo-Kibunja (Nkondi), S. Mũtugi Kĩgankah (Kajuki, Nithi Division) and Njeri Kĩnyua, all current or former students at the University of Nairobi. (Sub-locations are in Tharaka Division, unless otherwise specified.) In particular, I would like to thank S. Mũtugi Kĩgankah for tone-marking and Nyaga Mzalendo-Kibunja for rechecking the data. The research reported in this paper was funded by the Social Science Research Council, New York City, and conducted in Kenya under Research Permit #RDST/13/001/12C130/18, granted to me under the name Carolyn Harford Perez by the Office of the President, Republic of Kenya. I am grateful to Sue Warga of the SSRC, and Mrs. C.A. Mwangi and Miss L.A. Gwiyo of the Office of the President, Republic of Kenya, for all the help they have given me. I am also greatly indebted to members of the Department of Linguistics and African Languages at the University of Nairobi: Prof. M.H. Abdulaziz, Dr. Karega Mũtahi, Prof. Lucia Omondi, Dr. Okoth Okombo and Dr. Kithaka wa Mbeeria. I would also like to thank Corinne A. Kratz for all of her friendship and help in Nairobi. Thanks also to Katherine Demuth for providing greatly needed commentary and advice. Other comments have been provided by Alex Alsina, Victoria Bergvall, Joan Bresnan, Jonni Kanerva, 'Malillo Machobane, Sam Mchombo, Lioba Moshi, and Laurie Stowe; thanks again. I would also like to thank Robert Botne for meticulous editorial work. The usual disclaimers apply.

0. Introduction

Kitharaka, a Bantu language spoken in the Tharaka Division of Tharaka-Nithi District in Kenya, has a biclausal construction in which a verb with an athematic subject takes a finite complement. A referential NP in the athematic matrix subject position may receive its thematic role through linkage with any argument at a lower level of embedding. This construction will be referred to as Empty Operator Raising (EOR) in anticipation of an analysis which involves empty operator movement and because of the semantic similarity of EOR predicates to Raising to Subject (RS) predicates in English and other languages.

The paper consists of the following sections. Section 1 describes EOR. Sections 1.1 and 1.2 discuss, respectively, thematic transmission and clefting in EOR. Section 2 proposes an analysis of EOR. Section 2.1 compares EOR to constituent questions in Kitharaka, while section 2.2 compares it with WH-movement. Section 3 concludes.

1. EOR in Kitharaka

This section describes EOR in Kitharaka. This construction, like similar constructions in other Bantu languages [cf. Perez (Harford) 1985], takes a finite complement in both its non-“Raised” and “Raised” variants, as seen in examples (1a) and (1b).¹ A non-finite complement is possible only if “Raising” has taken place;² this variant will not be discussed in this paper.

¹ The examples in this paper are rendered using the standard orthography of Kimeru, the dialect cluster whose members are the closest relatives of Kitharaka, with one exception: I have borrowed the symbol “~” from Gikuyu to distinguish the seven vowels common to all of these languages. Although detailed phonetic work has yet to be done on Kitharaka, “i” may be tentatively identified as [e], “e” as [ɛ], “ū” as [o], “o” as [ɔ]. Also, “b” may be identified with [β], “g” with [ɣ], and “th” with [ð]. Tones are marked where available. Acute accent indicates high tone; grave accent low tone. Ungrammatical examples are not marked for tone.

The following abbreviations are used throughout:

CONT	continuous	OM	object marker	RPST	remote past
FV	verb final vowel	PASS	passive	RS	raising-to-subject
HAB	habitual	PRF	perfect	SM	subject marker
INF	infinitive	PR	predicative	ST	stative
LOC	locative	PST	past	VS	verb stem
NPST	near past	REC	reciprocal		

A slash in the glosses indicates fusion of morphemes in this otherwise agglutinative language. Numbers in all examples are noun class numbers and follow the numbering system for Bantu noun classes originated by Bleek [1862] and now referred to as the Bleek-Meinhof system.

² An example of a non-finite complement when “Raising” has taken place:

mù-ntù úyú n-á-íy-ìk-èèn-è í-kù-ìy-à m-bùrì
 1-person this.1 PR-SM.1-know-ST-REC/PRF-FV PR-1-steal-FV 10-goats
 ‘This person is known for stealing goats.’

- (1) a. *í-kw-iy-ìk-èèn-è* *àtĩ mù-ntù ù-yù n-áá-ìj-ìr-è* *m-bùrì* ³
 PR-SM.17-know-ST-REC/PRF-FV that 1-person this-1 PR-SM.1-steal-PST-FV 10-goats
 ‘It is known that this person stole goats.’
- b. *mù-ntù ù-yù n-áá-ìy-ìk-èèn-è* *àtĩ n-áá-ìj-ìr-è* *m-bùrì*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that PR-SM.1-steal-PST-FV 10-goats
 ‘This person is known to have stolen goats’

The embedded clauses in both examples (1a) and (1b) contain a tensed verb whose subject marker (SM) agrees obligatorily with the noun class of the matrix subject.⁴ In examples such as (1b), it is not possible to have distinct referential subjects in the matrix and embedded clauses. Thus, example (2), in which the subject *antũ* ‘people’ of the matrix verb is distinct from *muntũ ũra* ‘that person’, subject of the embedded clause, is ungrammatical.

- (2) **a-ntũ i-ba-iy-ìk-een-e* *aĩ mu-ntũ ũ-ra*
 2-people PR-SM.2-know-ST-REC/PRF-FV that 1-person 1-that
n-aa-raa-um-ir-e nja ya n-thĩ igoro
 PR-SM.1-NPST-leave-PST-FV outside of 9-country day before yesterday
 ‘People are known that that man left the country the day before yesterday’

There is an exception to this generalization. The matrix and embedded subjects may be distinct, but two conditions must be met, as listed in (3) and illustrated by the example in (4).

³ I refer to the first morpheme of both the matrix and embedded verbs as the predicator (PR), a term I owe to Patrick R. Bennett. Bergvall [1987] uses the term “assertion marker” for a cognate morpheme in Gikuyu. Clements [1985] refers to this morpheme as the focus particle. The PR takes the form *i-* preceding a consonant and *n-* preceding a vowel, with lengthening of the vowel. It appears on both verbs and nouns, with significant syntactic effects, as will be seen below.

The stative extension in examples (1a, b) is a verbal morpheme which intransitivizes a verb, in a way similar to the passive. In this verb, and in another in the examples in this paper, it is found in combination with the reciprocal extension, which carries no meaning in this context. This “spurious reciprocal” phenomenon, in which the reciprocal occurs meaninglessly with the stative, is also found in Kiswahili [Cf. Mchombo [1990] on the stative extension in Chichewa.]

⁴ Kitharaka is similar in many respects to Bantu languages already described in the generative literature (cf., for example, Bennett [1970], Bergvall [1987], Bokamba [1981], Bresnan and Mchombo [1987], Bresnan and Moshi [1990], Clements [1984], [1985], Demuth and Johnson [1989], Demuth and Gruber [1994], Perez (Harford) [1985]), and readers are referred to these sources and the references therein for further information. Other sources on this language include Mbeeria [1979, 1981] and Lindblom [1914].

- (3) a. The matrix subject must be linked to an embedded position other than the subject.
 b. There must be a clefted preverbal NP in the embedded clause which is linked to the embedded subject (represented by the subject marker on the verb).

- (4) *mù-ntù ù-yù n-áá-íy-ìk-èèn-è* *n-úú-rúáru*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV PR-14-illness
bù-tim-īt-è *à-ùùth-à*
 SM.14-cause-PRF-FVSM.1-become thin-FV

‘This person is known that it is an illness that made him/her thin.’

In example (4), the preverbal NP *nūūrūaru* ‘illness’ is clefted and linked to the subject marker *bū-* of *būtīmīte* ‘cause’, the causative verb, via class 14 agreement between the noun *ūrūaru* ‘illness’ and *bū-*. The matrix subject *muntū ūyū* ‘this person’ is linked to the subject of the verb *auutha* ‘became thin’, also via noun class agreement (class 1), in the clause embedded under *būtīmīte*. Example (5) shows that the clefted preverbal NP is obligatory.

- (5) **mu-ntū ū-yū n-aa-íy-ìk-èèn-e* *bū-tim-īt-e*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV SM.14-cause-PRF-FV
a-uuth-a
 SM.1-become thin-FV

‘This person is known that it made him thin.’

The subject marker of *būtīmīte* is class 14, which could refer to the class 14 noun *ūrūaru* ‘illness’, omitted here, as any preverbal subject NP may be in this language. However, the subject marker of a verb immediately embedded under an EOR verb may not receive such a pronominal interpretation, but must be coreferential with the matrix subject in the absence of a clefted preverbal NP.⁵ *Būtīmīte* must be linked to an overt NP; hence, example (5), which has omitted this NP, is ungrammatical.

Not only is this preverbal NP obligatory, but it must also be clefted. An embedded subject which is not clefted, such as *ūrūaru* ‘illness’ in example (6), is ungrammatical.

⁵ In this respect these data may be compared to the Kiruundi data presented in Perez (Harford) [1984], in which the same coreference requirement on the embedded subject holds for finite verbs embedded under Control verbs. They may also be compared to English examples such as *John wants Mary to leave* and *John wants to leave*, in which the absence of *Mary* forces the subject of the infinitive *to want* to be coreferential with *John*.

- (6) **mu-ntũ ũ-yũ n-aa-iy-ĩk-een-e* *ũ-rũaru bũ-tim-ĩt-e*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV 14-illness SM.14-cause-PRF-FV
a-uuth-a
 SM.1-become thin-FV

‘This person is known that an illness made him/her thin’

There are two other observations to be made about the EOR variant in which the matrix and embedded subjects are not linked to each other. First, the requirement for a clefted preverbal NP cannot be exclusively due to the change of subject between the matrix and embedded clauses, because it is possible to have an overt pronoun preceding the embedded verb which is linked to the matrix subject, as in (7a), as long as it is clefted. The example in (7b) is ungrammatical because the preverbal pronoun in the embedded clause is not clefted.

- (7) a. *mũ-ntũ ũ-yũ n-áá-iy-ĩk-èèn-è* *àtĩ n-wé à-ij-ir-è* *m-bũrĩ*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that PR-PRO.1 SM.1-steal-PST-FV 10-goats

‘This person is known that it is s/he who stole the goats.’

- b. **mu-ntũ ũ-yũ n-aa-iy-ĩk-een-e* *aĩ we a-ij-ir-e* *m-bũrĩ*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that PRO.1 SM.1-steal -PST-FV 10-goats

‘This person is known that s/he stole the goats.’

The only way for a bare NP to appear in preverbal position in the embedded clause is for it to be linked to a clefted pronoun marked with the PR, as illustrated in (8). In this example, the independent pronoun *bũ* is the clefted preverbal NP and *ũrũaru* is a topic linked to it.

- (8) *mũ-ntũ ũ-yũ n-áá-iy-ĩk-èèn-è* *ũ-rũárú*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV 14-illness
ĩ-bũ bũ-tim-ĩt-è *à-ùuth-à*
 PR-PRO.14 SM.14-cause-PRF-FV SM.1-become thin-FV

‘This person is known that an illness, it is it that made him/her sick.’

Second, it is also possible for the matrix subject to be coreferential with the object of the embedded clause rather than with a subject at a deeper level of embedding. As before, there must be a clefted preverbal NP. Examples (9a) and (9b) together show that non-clefting is ungrammatical. Examples (10a) and (10b) show that the object marker (OM) on the verb is optional and does not affect the requirement for a clefted preverbal NP.

- (9) a. ?*mù-ntù ù-yú n-áá-íy-ìk-èèn-è* *àñ í-cìngáù òdrag-ír-é*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that PR-Cingau SM.1-kill-PST-FV
 ‘This person is known that it is Cingau that killed.’
- b. **mu-ntù ù-yú n-aa-íy-ìk-èèn-è* *añ cingau oorag-ír-e*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that Cingau SM.1-kill-PST-FV
 ‘This person is known that Cingau killed.’
- (10) a. *mù-ntù ù-yú n-áá-íy-ìk-èèn-è* *àñ í-cìngáù à-mù-úrág-ír-é*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that PR-Cingau SM.1-OM.1-kill-PST-FV
 ‘This person is known that it is Cingau that killed him/her.’
- b. **mu-ntù ù-yú n-aa-íy-ìk-èèn-è* *añ cingau a-mù-úrág-ír-e*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that Cingau SM.1-OM.1-kill -PST-FV
 ‘ This person is known that Cingau killed him.’

Note the contrast with RS in a language like English, in which examples such as **This person seems that John killed (him)* and **This person seems John to have killed (him)* are impossible.⁶

To summarize at this point, in EOR, the matrix and embedded subjects are obligatorily coreferential, unless the matrix subject is linked to another embedded position and the embedded clause contains a clefted preverbal NP modified by a relative clause. The embedded subject is never interpreted as a pronoun.

1.1. Thematic Transmission. In the examples in the previous section, the matrix subject is not assigned a thematic role by the matrix verb, but receives one by being coreferential with an argument in the embedded clause. The evidence for thematic transmission in this construction comes from expletive subjects, idioms, and the obligatoriness of the embedded clause.

⁶ Examples in (9)-(10) do not cover all of the possibilities for the linking of the matrix subject in this construction, such as indirect object, possessor of an object, etc. There are two reasons why I prefer to confine the data to the possibilities presented here. First, the most important characteristic of this construction, on which the analysis to be presented below is based, is that the matrix subject need not be linked to the immediately embedded subject. The matrix subject must be linked to receive a thematic role, but the range of possible sources for this thematic role is tangential to the obligatory linking of the embedded subject. Second, the examples given already push the limits of the elicitation method, which requires reasonably natural and pragmatically plausible utterances to be fully reliable. Beyond a certain point, data become so artificial and pragmatically implausible that consultant judgements are unreliable. For these reasons, I will assume that there are no purely syntactic limitations on the linking of the matrix subject, since the analysis of this construction does not hinge on there being any.

1.1.1. Expletive Subjects. In example (1a), repeated here as (11), *muntū ūyū* appears as the subject of the embedded clause and the EOR verb has the Class 17 locative subject marker. The locative subject marker *kw-* is used here as an expletive, parallel to the use of English *there* in non-thematic positions. It may not be the subject marker of a verb with a thematic subject, as illustrated in (12).

- (11) *í-kw-iy-ìk-èèn -è* *àfì mù-ntū ū-yū n-áá-ìj-ir-è* *m-bùrì*
PR-SM.17-know-ST-REC/PRF-FV that 1-person this-1 PR-SM.1-steal-PST-FV 10-goats

‘It is known that this person stole goats.’

- (12) **i-kw-a-cam-ir-e* *ũ-cūrũ*
PR-SM.17-RPST-taste-past-FV 14-gruel

‘There tasted gruel.’

The possibility of an expletive subject indicates that the subject of this verb is athematic. The reader should consult Perez (Harford) [1983], Perez (Harford) [1985], and Bresnan and Kanerva [1988] for further discussion of the expletive use of locative morphology in Bantu languages.⁷

1.1.2. Idioms. Evidence that the matrix subject is thematically the subject of the embedded clause comes from idioms. The following sentence (13) means that an initiation candidate has completed the first stage of initiation. The idiomatic meaning is preserved when the subject of the idiom is the subject of *kūiyíkana*, as in (14).

- (13) *kī-rīmo i-kī-a-mer-ir-i-e*⁸ *mu-ntū ū-yū*
7-animal PR-SM.7-RPST-swallow-PST-VS-FV 1-person this-1

‘The “kirimo” swallowed this person.’

[= This person has completed the first stage of initiation.]

- (14) *kī-rīmo i-kī-iy-ìk-een-e* *afì i-kī-a-mer-ir-i-e*
7-animal PR-SM.7-know-ST-REC/PRF-FV that PR-SM.7-RPST-swallow-PST-VS-FV

mu-ntū ū-yū
1-person this-1

‘The “kirimo” is known to have swallowed this person.’

⁷ This pattern also occurs with the verbs *kùmènyèkànà* ‘to be known’ and *gũtìgwà* ‘to be thought’.

⁸ Certain morphemes, including the past morpheme *-ir-* and the habitual morpheme *-ag-* are infixed into verb stems ending in the vowel [i], as in this example and in examples (19a, b) that follow.

The observation that idiomatic meaning is preserved indicates that *kĩrĩmo* is the thematic subject of the embedded verb, not of the matrix verb *kũiyĩkana*.

1.1.3. Obligatory Presence of Embedded Clause. An apparent counterexample to the proposal that the thematic role of the matrix subject comes from the embedded clause is found in examples such as that in (15). The subject of this sentence must bear a thematic role, yet there is no embedded clause for this role to come from, indicating that the role comes from the matrix clause. This in turn suggests that the thematic role of the subject also comes from the matrix clause in the examples where there is an embedded clause. In that case, the matrix and embedded subjects would be bound without thematic role transmission.

- (15) *mu-ntũ ũ-yũ n-aa-iy-ĩk-een-e* *bweega mono*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV well very
 ‘This person is very well known.’

However, the verb in example (15) represents a distinct subcategorization for the same lexical item in whose predicate argument structure the subject bears a thematic role not derived by any thematic role assigning process. Unlike the verb which is followed by an embedded clause, this verb may not bear expletive concord, as shown by the ungrammatical example in (16).

- (16) **i-kw-iy-ĩk-een-e* *mu-ntũ ũ-yũ bweega mono*
 PR-SM.17-know-ST-REC/PRF-FV 1-person this-1 well very
 ‘There is known this person very well’

It is possible for a predicate whose subject thematic role has been suppressed by a process such as passivization to bear expletive concord, as in (17). Therefore, the example in (16) is not ungrammatical because of a failure of Case assignment following an intransitive verb. Rather, it is ungrammatical because the embedded clause is an obligatory constituent of the verb, which does not assign a thematic role to its subject. It is this verb that is the EOR verb.

- (17) *i-kw-aand-ag-w-a* *mp-eempe mũ-nda-ni ũ-ra*
 PR-SM.17-plant-HAB-PASS-FV 9-maize 3 -field-LOC 3-that
 ‘There is maize planted in that field.’

To summarize at this point, this section has developed arguments for thematic transmission in EOR on the basis of evidence from expletive subjects, idioms, and the obligatoriness of the embedded clause.

1.2. Clefting. This section presents evidence that a clefted NP is clefted by the addition of the assertion marker (PR), which removes it from the clause it is linked to.⁹ There are two kinds of evidence that the PR removes an NP from a clause. First, an NP marked with the PR appears at the left periphery of the clause, as in (18).

- (18) *i-ka-ĩjĩ ga-ka baaba a-gwat-ir-e ga-kĩ-iy-ag-a*
 PR-13-boy this-13 1.Father SM.1-catch-PST-FV SM.13-CONT-steal-HAB-FV
 ‘This boy, Father caught him stealing.’

Second, the clause following an NP marked with the PR is a relative clause.¹⁰ Evidence that this is the case comes from an alternation in the class 1 subject marker: it takes the form *a-* in main clauses and object relative clauses, but may appear as *ũ-* in subject relative clauses (*w-* before vowels), as illustrated by the examples in (19).¹¹ This alternation is not possible in non-relatives (20).

- (19) a. *ĩ-mb-ón-ir-è mù-ntũ ú-rá w-èènd-ág-ì-à mà-úà.*
 PR-I-see-PST-FV 1-person 1-that REL.SM.1-sell-HAB-VS-FV 6-flowers
 ‘I saw the person who sells flowers.’
- b. *n-w-ĩjĩ mù-ntũ ú-rá èènd-ág-ì-à mà-úà?*
 PR-you-know 1-person 1-that SM.1/sell-HAB-VS-FV 6-flowers
 ‘Do you know the person who sells flowers?’

- (20) a. *màitũ òòn-ir-e bààbà* ‘Mother saw Father.’
 1.Mother SM.1/see-PST-FV 1.Father
- b. **maitũ w-on-ir-e baaba*
 1.Mother REL.SM.1-see-PST-FV 1.Father

⁹ This being one of several functions of the PR. Note that this assumption applies only to NPs marked with the PR, not to verbs.

¹⁰ Other Bantu languages in which clefts are followed by relative clauses include Chichewa [Bresnan & Mchombo 1987], Kihung’an [Takizala 1972], Gikuyu [Clements 1985] and Kiruundi [Sabimana 1986].

¹¹ This alternation is optional for my consultants, all of whom are younger speakers. I suspect that it is obligatory for older speakers. It is presented as obligatory in descriptions of Gikuyu. Note also that two other potential types of evidence for relative clausehood in the Thaagicu subgroup and in Bantu languages in general are not available for Kitharaka. There is no tonal evidence that I know of; that is, no evidence of an initial high tone, although not enough work has been done on the tonal system of this language to be sure. Also, the distinction between the *-ti-* and *-ta-* negative markers, which may be used to distinguish relative from non-relative clauses in Gikuyu [Barlow 1951, Bergvall 1987, Clements 1984], may not be used in Kitharaka, since *-ti-* and *-ta-* are attested in both relatives and non-relatives.

The alternate form is possible in clauses following NPs marked with the PR, providing evidence of their relative status, as in (21). The example in (22) provides a sample of its occurrence in EOR.

- (21) *i-mu-ntũ ũ-yũ ũ-raa-ij-ir-e m-bũri*
 PR-1-person this-1 REL.SM.1-NPST-steal-PST-FV 10-goats
 ‘It is this person who stole goats.’

- (22) *mũ-ntũ ú-yũ n-áá-iy-ìk-èèn-è àtĩ n-wé*
 1-person this-1 PR.SM.1-know-ST-REC/PRF-FV that PR-PRO.1
ũ-ráá-ij-ir-è m-bũrì
 REL.SM.1-NPST-steal-PST-FV 10-goats
 ‘This person is known that it is s/he that stole goats.’

A second property of relative verbs, unrelated to noun class, is that a relative verb may not bear the PR (23), nor may a verb in a clause following an NP which is marked with the PR (24). Significantly, the EOR case (25), as well, does not permit PR.

- (23) **i-mb-on-ir-e mu-ntũ ũ-ra n-w-eend-ag-i-a ma-ũa*
 PR-I-see-PST-FV 1-person 1-that PR-REL.SM.1-sell-HAB-VS-FV 6-flowers
 ‘I saw the person who sells flowers.’

- (24) **i-ka-ĩjĩ ga-ka baaba n-aa-gwat-ir-e ga-kĩ-iy-ag-a*
 PR-13-boy this-13 1.Father PR.SM.1-catch-PST-FV SM.13-CONT-steal-HAB-FV
 ‘This boy, Father caught him stealing.’

- (25) **mu-ntũ ũ-yũ n-aa-iy-ìk-èèn-e n-ũũ-rũaru*
 1-person this-1 PR.SM.1-know-ST-REC/PRF-FV PR-14-illness
i-bũ-tim-ĩit-e a-uuth-a
 PR-SM.14-cause-PRF-FV SM.1-become thin-FV
 ‘This person is known that it is his/her illness which made him/her thin.’

To summarize, an NP is clefted by the PR, which removes it from its clause. Two kinds of evidence for this removal are: 1) the clefted NP appears at the left periphery of the clause; 2) the clefted NP is followed by a relative clause. Evidence that this clause is a relative comes from the observations that it may take the class 1 relative subject marker *ũ-* and that it may not bear the predicator marker. Both of these types of evidence indicate that the clefted preverbal NP in the embedded clause of a EOR verb has been removed from its clause.

2. Analysis of EOR

In this section, I propose an analysis of EOR using the Principles and Parameters (P & P) framework [Chomsky 1981, Haegeman 1994]. The properties of the EOR construction illustrated in the previous section may be summarized as in (26).

- (26) Properties of the EOR construction
- a. Its matrix subject position is athematic. (§ 1.1)
 - b. Its embedded complement may be finite or non-finite. (examples (1a-b))
 - c. A “raised” matrix subject corefers with the subject or a non-subject of an embedded finite clause. (examples (1b), (9a), (10a))
 - d. If there is a subject NP in the embedded clause, it must be clefted. (examples (4), (5))
 - e. If there is no subject NP in the embedded clause, the matrix subject corefers with the empty embedded subject. (examples (4), (6))

I assume that properties (26a-b) are stipulated as part of the lexical entries of EOR verbs. Property (26c) raises the question of how coreference is established between the matrix subject and a position in the embedded clause. Pursuing a familiar strategy, one could propose that the matrix subject is moved to the matrix clause from the embedded clause, establishing coreference through the linkage between the two positions. This proposal encounters two problems.

First, interclausal movement is ruled out in this context by several principles of P & P theory. The theory recognizes two types of movement: WH-movement and NP-movement. WH-movement between the two clauses of an EOR construction is ruled out by the reasonable assumption that the matrix subject occupies an A-position (an argument position), not an A'-position (a non-argument position). WH-movement takes place only to an A'-position. NP-movement is also ruled out, by two separate requirements which interact to restrict NP-movement to the subjects of non-finite clauses. Movement between finite clauses violates Principle A of the Binding Theory. In addition, the subject and object positions of finite clauses are both Case-marked positions, and traces of NP-movement are required to be Caseless.

Second, the proposal sheds no light on properties (26d-e), the properties of EOR which crucially distinguish it from constituent question formation, which may be straightforwardly analyzed as involving WH-movement (see §2.2). There is no way I know of for either type of movement to force the clefting of an intervening subject NP.

However, the theory provides another way of accounting for coreference which does not involve interclausal movement and which also offers a way to account for properties (26d-e). Consider the Tough Movement construction in English (27).

(27) *Mary is tough to cheat.*

The meaning of this sentence entails that *Mary* is the object of the verb *cheat*, but it does not occupy the canonical object position, the position immediately following *cheat*, to which it assigns a theta role and Case. Instead, in accordance with the Projection Principle [Chomsky 1981], this position is occupied by an empty category, which must be linked to *Mary* to establish coreference. However, for reasons just discussed, this linkage cannot be established directly through interclausal movement. Instead, a non-overt element, referred to as an empty operator, is base-generated in the object position in the embedded clause and moved to the immediately embedded SPEC/CP, following standard assumptions about empty operator movement (Chomsky [1977], Contreras [1993], Lasnik and Stowell [1991]); for a contrary view about the landing site of the empty operator, see Authier [1989]. The matrix subject is base-generated in its surface position and coindexed with the empty operator in SPEC/CP, establishing coreference. In this way, the idea of accounting for coreference by movement is retained, but in a way that avoids violating theoretical principles that are otherwise well-established.¹²

These ideas may be extended to EOR in Kitharaka. Recall example (1b), repeated here as (28). In this example, I propose that an empty operator (\emptyset_i) has been base-generated in the embedded subject position, satisfying the Extended Projection Principle [Chomsky 1981], which requires all sentences to have subjects. It has then been moved to the embedded SPEC/CP, leaving behind a coindexed trace (t_i). The empty operator is coindexed with the matrix subject, which is linked in this way to the embedded subject and accordingly inherits its theta role. The example in (28) may then be represented as in (29).

(28) *mù-ntũ ú-yũ n-áá-iy-ìk-èèn-è àñ n-áá-ij-ìr-è m-bũrì*
 1-person this-1 PR-SM.1-know-ST-REC/PRF-FV that PR-SM.1-steal-PST-FV 10-goats
 ‘This person is known to have stolen goats’

(29) Muntũ ũyũ naiyĩkeene ati [\emptyset_i [t_i na_iijire mbũri]]
 ‘This person is known to have stolen goats.’

¹² Other constructions for which empty operators have been proposed include *that*-relatives and parasitic gaps in English. See Haegeman [1994] for a discussion of empty operators in *that*-relatives.

As the representation indicates, the matrix subject is linked to the embedded subject via the empty operator without interclausal movement. The case in which the matrix subject is linked to an embedded non-subject receives the same analysis (30).

- (30) Muntū ūyū_i; naiyīkeene aŋi [\emptyset _i [nūūrūaru_j [t_i [e_j bŭ_jtimūite [t_i a_iuutha]]]]
-

‘This person is known that it is an illness that made him/her thin.’

The empty operator also provides an account of properties (26d-e), the constraints on the immediately embedded subject. The property which is the key to the analysis is (26d), the obligatory clefting of an embedded subject NP which is not coreferential with the matrix subject. Why is clefting of such a subject obligatory? Recall from section 1.2 that the effect of clefting is to remove an NP from its clause. Why is removal of this NP obligatory in this context? One possibility, suggested by the work of Ouhalla [1993] is that it must be removed to maintain a certain distance from something else. What is this something else? The most obvious answer is that it is the empty operator in SPEC/CP. Setting aside examples (7a-b) for the time being, the following descriptive generalization emerges from the rest of the data.

- (31) If there is no agreement between the matrix and embedded subjects, the empty operator must not occur in the same CP as the immediately embedded subject.

This requirement on the empty operator—that it maintain a certain distance from the immediately embedded subject—is reminiscent of Ouhalla’s [1993] Anti-Agreement Effect (AAE). The AAE describes a phenomenon in which local extraction of a subject forces neutral agreement on the verb of the clause it was extracted from.¹³ Ouhalla’s idea (see also Borer [1984]) is that a WH-operator in SPEC/CP may bind the immediately following subject position, creating unacceptability under certain conditions. Without adopting Ouhalla’s explanation of the AAE (rich subject AGR licenses a resumptive pronoun in subject position, which then cannot be bound by the WH-operator), the basic proposal here for EOR is that the clefting of an embedded subject NP distinct in reference from the matrix subject is a strategy to create enough distance between the empty operator and the following subject position to prevent the former from binding the latter. This strategy is required precisely when there is a change in subject because the empty

¹³ The AAE exists in Kitharaka, as illustrated in example (19a) above.

operator and the embedded subject bear different indices and therefore cannot form a chain.

For example, compare (1b) and (5), represented schematically in (32) and (33), respectively. In (32), the SM of the embedded verb agrees in noun class with its subject, the trace of the empty operator, which, in turn, is coindexed with the matrix subject. Therefore, the three may form a chain, and the empty operator acceptably binds the embedded subject position. In (33), on the other hand, the SM of the embedded verb belongs to a noun class different from that of the matrix subject. Therefore, the empty operator cannot bind the embedded subject without creating a chain with different indices, and the example is ungrammatical.

(32) Muntū ūyū_i naiyīkeene afi [\emptyset_i [t_i na_iijire mbūri]]

‘This person is known to have stolen goats.’

(33) *Muntū ūyū_i naiyīkeene afi [\emptyset_i [e_j bū_jtimĩite [t_i a_iuutha]]

‘This person is known that it made him/her thin.’

Now compare (4) and (6), represented here as (34) and (35), respectively. The example in (35) is ungrammatical for the same reason as that in (33): it binds the embedded subject *ūrūaru*, which has a distinct index, as well as being an R-expression, which cannot be bound in any case, according to Condition C of the Binding Theory. In (34), the clefting of *ūrūaru* places the nearest embedded subject position in a relative clause modifying the focus of the cleft. Assuming that the empty operator is still in the highest SPEC/CP, it is no longer close enough to the nearest subject position to bind it.

(34) Muntū ūyū_i naiyīkeene afi [\emptyset_i [nūrūaru_j [t_i [e_j bū_jtimĩite [t_i a_iuutha]]

‘This person is known that it is an illness that made him/her thin.’

(35) *Muntū ūyū_i naiyīkeene afi [\emptyset_i [ūrūaru_j bū_jtimĩite [t_i a_iuutha]]

‘This person is known that an illness made him/her thin.’

Now consider (7a) and (7b), represented here in (36) and (37), respectively. The same pattern holds that was observed earlier: the overt pronoun *we* ‘s/he’ cannot occur in the embedded clause without being clefted. However, since the pronoun is

coreferential with the matrix subject, (37) cannot be ungrammatical because of the need to avoid an illicit chain.

(36) Muntū ūyū_i naiyīkeene afi [\emptyset _i [nwe_i [t_i a_iijire mbūri]]]

‘This person is known to have stolen goats.’

(37) *Muntū ūyū_i naiyīkeene afi [\emptyset _i [we_i a_iijire mbūri]]

‘This person is known that s/he who stole goats.’

I propose that the ungrammaticality of (37) is similar to that of (35), which involves a Condition C violation as well as an illicit chain. In this instance, the problem with (37) is that the empty operator binds the independent pronoun *we*, a violation of Ouhalla’s formulation of Aoun and Li’s A'-disjointedness Requirement, stated in (38) (See also Aoun and Hornstein [1986], Aoun and Li [1993].

(38) A pronoun must be free from the most local A'-binder in the smallest C(omplete) F(unctional) C(omplex) which contains the pronoun. [Ouhalla 1993:506]

The smallest CFC is the minimal CP which contains the pronoun (*ibid.*, p.490).

To summarize, the obligatory clefting of an NP distinct from the matrix subject in an EOR construction has been argued to be a strategy to distance the embedded subject position from the EO in the immediately embedded SPEC/CP, to avoid forming an illicit chain. This strategy is also required when an independent pronoun occurs in the embedded clause, to avoid violating the A'-disjointedness Requirement.

2.2. EOR and WH-Movement This section compares EOR with WH-movement in Kitharaka. The analysis in the preceding section is based on the idea that an empty operator in SPEC/CP binds the following subject position. However, not all exhibit the same behavior. For example, relative clauses, which can reasonably be assumed to be WH-constructions in Kitharaka, fail to show the patterns of coreference and clefting illustrated above for EOR. The following are examples of constituent questions, in which the interrogative is clefted and placed at the left periphery. As seen above in section 1.2, clefted NPs are followed by relative clauses.

As the examples in (39) show, the question word *nūū* may be linked to either the subject (39a) or the object (39b-c) of the embedded clause. Note, however, the status of the subject in the examples in which the object has been extracted. There is no preverbal NP in the embedded clause in (39b), but there is no obligatory link

between the embedded subject and the extracted wh-word. Contrast this example with the examples in (33), (34), and (35) above, which show that in EOR, the embedded subject must be bound by the matrix subject unless it is a clefted preverbal NP. Note also that in example (39c), there is a preverbal NP in the embedded clause, but it is not clefted.

(39) a. *n-ũũ maitũ a-ug-ir-e afi n-oon-ir-e baaba*
 PR-who 1.Mother SM.1-say-PST-FV that PR-SM.1-see-PST-FV 1.Father
mũ-cemanio-ni?
 3-meeting-LOC

‘Who did Mother say saw Father at the meeting?’

b. *n-ũũ maitũ a-ug-ir-e afi n-oon-ir-e mũ-cemanio-ni?*
 PR-who 1.Mother SM.1-say-PST-FV that PR-SM.1-see-PST-FV 3-meeting-LOC

‘Who did Mother say s/he saw at the meeting?’

c. *n-ũũ maitũ a-ug-ir-e afi baaba n-oon-ir-e mũ-cemanio-ni?*
 PR-who 1.Mother SM.1-say-PST-FV that 1.Father PR-SM.1-see-PST-FV 3-meeting-LOC

‘Who did Mother say Father saw at the meeting?’

Similar differences in the binding properties of WH-operators have been observed in other languages, particularly by Ouhalla [1993], who points out that in languages such as Fiorentino and Trentino, intermediate traces trigger the AAE, whereas in languages like Berber and Breton, they don’t (p.491). Also, differences in the type of WH-operator (QP vs. empty operator) have been correlated with crossover effects in English by Lasnik and Stowell [1991]. An explanation of the split in binding effects found in EOR and constituent questions in Kitharaka may be pursued along either of these lines, but I will leave the issue open.

3. Conclusion

This paper has described a biclausal construction in Kitharaka, referred to as Empty Operator Movement (EOR), which combines an athematic matrix subject with a finite complement. “Raising” to the matrix subject position may take place from a subject or a non-subject position in this finite clause. The assumption that empty operator movement is possible in finite clauses in this language has been used to account for a syntactic property not found in any English construction, namely the requirement that if the subject of the immediately embedded clause is distinct from the matrix subject, there must be a clefted NP in the embedded clause. This requirement has been attributed to the need to avoid the binding of the embedded subject by the empty operator in the immediately embedded SPEC/CP.

REFERENCES

- Aoun, Joseph and Norbert Hornstein. 1986. "Bound and Referential Pronouns" To appear in James Huang and Robert May (eds.), *Logical Structure and Linguistic Structure*. Dordrecht: Kluwer.
- Aoun, Joseph and Audrey Li. 1993. *Syntax of Scope*. Cambridge, MA: MIT Press.
- Authier, J.-Marc. 1989. "Two Types of Empty Operator." *Linguistic Inquiry* 20: 117-25.
- Barlow, A. Ruffell. 1951. *Studies in Kikuyu Grammar and Idiom*. Edinburgh: William Blackwood and Sons.
- Bennett, Patrick R. 1970. "The Problem of Class in Kikuyu." *African Language Studies* 8:48-59.
- Bergvall, Victoria L. 1987. "Focus in Kikuyu and Universal Grammar." Ph.D. dissertation, Harvard University.
- Bleek, W.H.I. 1862. *A Comparative Grammar of South African Languages*. Cape Town: J.C. Juta.
- Bokamba, Eyamba. 1981. *Aspects of Bantu Syntax*. Preliminary edition, Department of Linguistics, University of Illinois, Champaign-Urbana.
- Borer, Hagit. 1984. "Restrictive Relatives in Modern Hebrew." *Natural Language and Linguistic Theory* 2:219-60.
- Bresnan, Joan and Jonni Kanerva. 1988. "Locative Inversion in Chicheŵa: A Case Study of Factorization in Grammar." *Linguistic Inquiry* 20:1-50.
- Bresnan, Joan and Sam Mchombo. 1987. "Topic, Pronoun and Agreement in Chicheŵa." *Language* 63:741-782.
- Bresnan, Joan and Lioba Moshi. 1990. "Object Asymmetries in Comparative Bantu Syntax." *Linguistic Inquiry* 21:147-185.
- Chomsky, Noam. 1977. "On *wh*-movement." In P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax*, pp. 71-133. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

- Clements, George N. 1984. "Principles of Tone Assignment in Kikuyu." In George Clements and John Goldsmith (eds.), *Autosegmental Studies in Bantu Tone*, pp. 281-339. Dordrecht: Foris.
- Clements, George N. 1985. "Binding Domains in Kikuyu." *Studies in the Linguistic Sciences* 14:37-56.
- Contreras, Helas. 1993. "On Null Operator Structures." *Natural Language and Linguistic Theory* 11:1-30.
- Demuth, Katherine and Jeffrey Gruber. 1995. "Constraining XP Sequences." In Victor Manfredi and Karl Reynolds (eds.), *Niger-Congo Syntax and Semantics* No. 6:3-30. Boston, MA: Boston University African Studies Center.
- Demuth, Katherine and Mark Johnson. 1989. "Interaction between Discourse Functions and Agreement in Setawana," *Journal of African Languages and Linguistics* 11:21-35.
- Haegeman, Liliane. 1994. *Introduction to Government and Binding Theory* (2nd edition). Oxford: Basil Blackwell.
- Lasnik, Howard and Tim Stowell. 1991. "Weakest Crossover." *Linguistic Inquiry* 22:687-720.
- Lindblom, G. 1914. *Outline of a Tharaka Grammar*. Uppsala: J.A. Lundell.
- Mbeeria, Kithaka wa. 1979. "The Morphology of Kitharaka Nominal Word." B.A. thesis, University of Nairobi.
- Mbeeria, Kithaka wa. 1981. "The Consonants of Kitharaka." M.A. thesis, University of Nairobi.
- Mchombo, Sam. 1990. "A Formal Analysis of the Stative Construction in Bantu." Paper presented at the 21st Annual Conference on African Linguistics, University of Georgia, Athens.
- Ouhalla, Jamal. 1993. "Subject-Extraction, Negation and the Anti-Agreement Effect." *Natural Language and Linguistic Theory* , 11:477-518.
- Perez (Harford), Carolyn. 1983. "Locative Pseudo-Subjects in Shona." *Journal of African Languages and Linguistics* 5:131-155.

Perez (Harford), Carolyn. 1984. "Anaphoric Binding in Bantu." In J. Drogo, V. Mishra and D. Testen (eds.), *Papers from the Twentieth Regional Meeting of the Chicago Linguistic Society* 20: 301-308.

Perez (Harford), Carolyn. 1985. "Aspects of Complementation in Three Bantu Languages." Ph.D dissertation, University of Wisconsin-Madison. [Distributed by the Indiana University Linguistics Club, Bloomington, IN, 1986.]

Sabimana, Firmard. 1986. "The Relational Structure of the Kirundi Verb." Ph.D dissertation, Indiana University.

Takizala, Alexis. 1972. "Focus and Relativization: The Case of Kihung'an" In John P. Kimball (ed.), *Syntax and Semantics*, Vol. 2, pp. 123-148. New York: Academic Press.

Department of Linguistics
University of Zimbabwe
Mt. Pleasant, Harare, Zimbabwe
harford@samara.co.zw

[Received September 1996;
revision received August 1997;
accepted September 1997]

