AGREEMENT-LESS DECLARATIVES IN NAIROBI SWAHILI

KAMIL UD DEEN
Kamil@hawaii.edu

The Swahili verbal complex is comprised of a set of discrete inflectional affixes that obligatorily occur in the same order and in all declarative indicative clauses (Ashton, 1947; Vitale, 1981; Krifka, 1995). The minimal verbal complex is: Subject Agreement – Tense – Verb Root – Mood. Subject Agreement in Swahili is sufficiently 'rich' (in the sense of Rizzi, 1982) to identify null subjects, and thus Swahili is a null subject language. In this paper I report on a phenomenon in Nairobi Swahili in which Subject Agreement can be omitted. I show that while speakers of Nairobi Swahili allow the omission of subject agreement infrequently (in approximately 5% of all indicative clauses), it nevertheless is grammatical. This raises the question of whether Nairobi Swahili allows null subjects in the absence of subject agreement. I show that null subjects indeed do occur in the absence of agreement, a fact that is puzzling given that there is no obvious identifier for the null subject. I propose an analysis in which a null constant occurs in subject position in the absence of agreement (Rizzi, 1992; 1997). This null constant is bound by an optionally null topic operator which surfaces in preverbal position. Thus the 'subjects' that occur in the absence of agreement are in fact topics. The syntactic restrictions on agreement omission support this analysis: in the absence of agreement, subjects cannot be quantifiers; an agreement-less clause cannot be used to answer a wh-question; an agreement-less clause cannot occur in embedded context. The results suggest that agreement omission is a process that is grammatical in language under restricted syntactic and discourse conditions.

1. INTRODUCTION

The omission of subjects in both adult and child language has received much attention in the literature in recent decades. It has been observed that languages largely fall into two categories: those that allow the subject argument to be omitted in matrix declarative clauses and those that require overt subjects. It has also been noted that this property correlates to a large degree with the richness of subject-verb agreement exhibited in any particular language (Taraldsen, 1978). Thus, English does not have rich subject verb agreement morphology and does not allow null subjects, while Italian has rich subject verb morphology and correspondingly allows null subjects.

(1)   a.   John watches Indy films for fun   English
         b.   * [e] watches Indy films for fun.

(2)   a.   Gianni mangia una mela.   Italian
         b.   [e] mangia una mela.

Let us call this the *pro* Generalization:

The *pro* Generalization: null arguments are permitted when corresponding rich agreement is available.

Rizzi (1982; 1986) argued that rich agreement allows the null element to receive identification. He argued that the subject position in such sentences is occupied by *pro*, which requires licensing and identification. Licensing occurs as it does with lexical subjects, i.e., through case marking. Because *pro* subjects are silent, they require identification, which occurs through rich agreement.

This generalization has been noted in the omission of other arguments, i.e., object drop is permissible when the language has rich object-verb agreement (see Huang, 1989). Additionally, Hebrew is a language that has rich agreement in the past tense, but not in the future tense. As expected under the generalization above, Hebrew allows null subjects only in the past tense, not in the present tense. Finally, Huang (1986; 1989) showed that an additional set of languages allow argument omission in the complete absence of any agreement whatsoever, e.g., Chinese, in which case identification occurs through discourse.

This phenomenon has generated a great deal of research and interest, and continues to be the subject of great study. The research that addresses this question assumes that the argument may/may not be omitted in the presence/absence of agreement. In this paper, I investigate this question from the opposite direction. I investigate a phenomenon in Nairobi Swahili in which the agreement morpheme itself is omitted. I investigate the frequency of this omission and the grammaticality of such omissions. Since Nairobi Swahili (like Standard Swahili) is a null subject language, I investigate whether the omission of the rich agreement affects the null-subject properties of Swahili. I conclude that omission of subject agreement occurs in topicalized structures in which the subject position is occupied by a null constant (Rizzi, 1992). This null constant is bound by the topic, and does not license agreement morphology.

2. SWAHILI MORPHOSYNTAX

In this section I will describe the inflectional affixes in Swahili, the nature of subject agreement, and the status of null subjects in Swahili.

2.1. *Inflectional Affixes*

Swahili is an agglutinative language, with considerable prefixing and suffixing. The unmarked word order is S-V-O, as shown in example (3)[1] below. In (1), the subject (Juma) occurs preverbally and the object (Mariam) occurs postverbally. The verb is embedded in a verbal complex which consists of subject agreement (*a-*) on the left periphery, followed by tense (*-na-*), object agreement (syllabic *–m-*) and then the verb root itself (*pend-*). The verb is followed by (in this case) one suffix which indicates mood (in this case indicative *–a*). The subject can be optionally absent (shown in example 4), and the identity of the subject is recoverable from the rich subject verb agreement. The subject may occur in postverbal position (5), with an obligatory pause and lower intonation (so-called comma intonation). Furthermore, the object may also be dropped (6).

	Subject	Verbal Complex	Object

(3) Juma   a - na - m – pend - a   Mariam
       Juma SA_{3s}-PRES- OA_{3s}- like - IND   Mariam
       'Juma likes Mariam'

(4)   A - na - m – pend - a   Mariam
       SA_{3s} - PRES- OA_{3s}- like- IND   Mariam
       'He likes Mariam'

(5)   ni - na - m – pend - a   Mariam, mimi
       SA_{1s}- PRES- OA_{3s}- like - IND   Mariam $Spro_{1s}$
       'I like Mariam'

(6)   a - na - m – pend - a
       SA_{3s}- PRES - OA_{3s} – like - IND
       'He likes her'

---

[1] SA=subject agreement, OA=object agreement, pres=present tense, pst=past tense, fut=future tense, IND=indicative mood, SUBJ=subjunctive mood, 1s=1st person, singular, 1pl= 1st person plural, etc. Because noun classes are not the focus of this paper, I do not gloss them in my examples so as to avoid complications in the glosses.

Below are the SA and OA paradigms, followed by the more common tense markers in Swahili.

Table 1. SA Paradigm

SA marker	Person / Number
ni	1st singular
u	2nd singular
a	3rd singular
tu	1st plural
mu	2nd plural
wa	3rd plural

Table 2. OA Paradigm

OA marker	Person / Number
ni	1st singular
ku	2nd singular
m	3rd singular
tu	1st plural
mu	2nd plural
wa	3rd plural

Table 3. Common Swahili Tense Markers

Tense Marker	Tense
na / a	Present
li	Past
ta	Future
me	Pres. Perfect

While I have described Swahili as an S-V-O language, there is a considerable amount of material that intervenes between the subject and the verb root, and the object and the verb root. However, as (5) above shows, when the subject is moved, all elements of the verbal complex (including Subject Agreement) remain with the verb. Similarly, if the object is preposed, as in (7) below, all elements of the verbal complex remain within the verbal complex in their original positions, including the object agreement marker. These examples show that the verbal complex behaves as a unit in Swahili.

(7)   Mariam, Juma   a - na - m – pend - a
         Mariam, Juma SA_{3s}- PRES - OA_{3s} – like - IND
         'Mariam, Juma likes [t]'

2.2 *Agreement versus Pronominal Clitic*

In the Swahili literature (and indeed in the wider Bantu literature) there is considerable discussion on the status of agreement markers in modern Swahili. The discussion centers on whether SA is agreement between the subject and the verb or whether it is a subject pronoun. In the former case, SA is agreement (in pre-minimalist terms, it is the

head of AgrSP) and the subject is a true subject (i.e., a DP in the specifier of AgrSP). In the latter case, SA is the DP in the specifier of AgrSP, and the preverbal DP is in a higher topic position. Keach (1995) discusses the characteristics of SA in Swahili, concluding that it is ambiguous between being a pronominal clitic and an agreement marker. However, in Deen (2002) I argue for an agreement analysis. In addition to agreement arguments from Keach (1995), I argue that if the preverbal DP is a topic, it should have the well-known properties of topics that have been established from work in other languages. One property of topics is that they cannot be quantifiers (Lasnik & Stowell, 1991; Rizzi, 1993):

(8) a. I did everything
  b. *Everything, I did (it)

(9) a. Nothing is impossible
  b. *Nothing, (it) is impossible

   In Swahili, this restriction also holds. In (10a), the object (*kila kitu*) is in object position, and is ungrammatical when topicalized, as in (10b).

(10) a. a – li –nunu – a   kila kitabu
    SA_{3s} –past–buy–IND   every book
     'She bought every book'

  b. * [kila kitabu]$_i$, a – li – (ki)–nunu– a   [t$_i$]
     every thing$_i$ SA_{3s}–PAST–(OA_7)–buy–IND [t$_i$]
     'Every book, she bought'

   Thus the restriction on quantified topics holds in Swahili. Under a pronominal analysis of SA, the preverbal DP is in topic position, and so a quantifier should be ungrammatical. However, as (11) shows, quantifiers are possible in preverbal position, suggesting that the preverbal DP is in subject position.

(11) a. kila mtoto   a – li – nunu– a ki – tabu
     every child SA_{3s}– PAST –buy–IND 7–book
     'Every child bought a book.'

  b. kila ki–tabu   ki – li – nunuli – w – a   na mtoto
     every 7–book SA_7– PAST – buy –passive–IND by child
     'Every book was bought by a child.'

Furthermore, the answer to a question cannot be a topic:

(12)   Who arrived early?
          - John arrived early
          - ?? As for John, he arrived early

In Swahili, the preverbal DP can be the answer to a question:[2]

(13)   a.   nani   a – li – fik – a   mapema
                   who SA_{3s}– PAST –arrive–IND early
                   'Who arrived early?'

         b.   ?? Juma,   a – li – fik – a   mapema
                      Juma, SA_{3s}– PAST –arrive–IND early
                      'Juma, he arrived early.'

         c.   Juma   a – li – fik – a   mapema
                   Juma SA_{3s}– PAST –arrive–IND early
                   'Juma arrived early.'

When the preverbal DP is topicalized (indicated by 'comma' intonation) as in example (13b)), it is awkward as an answer to the question in (13a). Thus topics cannot be the answer to questions in Swahili. In example (13c) with normal sentence intonation, the non-topicalized preverbal DP is grammatical as the answer to the wh-question in (13a). This supports the view that the preverbal DP (without 'comma' intonation) is not in topic position, but rather in subject position.

A final argument in favor of an agreement analysis comes from typology. One criterion that distinguishes clitic pronouns from agreement affixes is the freedom of word order: pronouns are generally more free to move relative to the verb, or allow the verb to move around the clitic. For example, Tagalog has a series of clitics, all of which are constrained by a second-position rule (Schachter, 1995, p.1425). The verb can precede the clitic or follow it, as can other words in the sentence, with the only restriction being that the clitic must be in second position. Affixes, on the other hand, must generally remain proximal to the verb, in the same structural configuration, and with the same set of (usually) inflectional elements between it and the verb. For example, languages in the Takic family (a Southern

---

[2] Thanks to Stan Dubinsky and Ivano Caponigro for discussions on this point.

California branch of Uto-Aztecan) have a subject marker that, similar to Swahili, is the focus of debate. Among the languages of the Takic family, SA has been particularly well-studied in four languages: Luiseño, Cupeño, Serrano, and Cahuilla. In Luiseño, the unmarked word order is shown in (14a) (examples are from Steele, 1995), where the clitic (*up*) is in second position following the subject (*hengeemal*):

(14)   a.   hengeemal up   heyiq   Subject-clitic-verb
            boy   3SG   is:digging
            'The boy is digging'

       b.   heyiq   up hengeemal   Verb-clitic-subject
            is:digging 3 SG boy
            'The boy is digging'

       c.   * hengeemal heyiq   up   Subject-verb-clitic
             boy   is:digging 3 SG

In (14a), the unmarked order is subject-clitic-verb. According to Steele (1995, p.1227), (14b) with the verb preceding the clitic is semantically non-distinct from (14a). (14c) – where the clitic sequence is not second – is ungrammatical. This is also true of two of the other three most well-studied languages: Cupeño and Serrano. Thus the order of the clitic and verb is free, provided the clitic is in second position. However, the fourth well-studied Takic language, Cahuilla, has a set of bound pronominal elements that are obligatorily preverbal. Thus the order clitic-verb is grammatical, but verb-clitic is ungrammatical irrespective of whether the clitic is in second position or not. These clitics are "generally taken to be prefixes rather than (pronouns)" (Steele, 1995, p.1227).[3] In making this distinction, Steele (along with Jacobs, 1975; Steele, 1977; Langacker, 1977) uses word order as a diagnostic for whether a subject marker is an agreement affix or a pronominal clitic, with the former being fixed in position with respect to the verb, and the latter being somewhat freer. See also

---

[3] The only examples Steele gives are designed to illustrate that these prefixes combine subject and object marking, and not to illustrate the unacceptability of free word order. One example is her example (7a):
     'echem-némiwe
     1pl/2SG-chased
     'We chased you.'
The fact that Steele glosses this example with a hyphen between the prefix and verb suggests that it behaves as a single unit, akin to the Swahili verbal complex, and her description of the facts suggests the same.

Zwicky & Pullum (1983) who use similar criteria for distinguishing between clitics and affixes.

We can now apply this test to the Swahili SA marker to determine whether it is a prefix or a pronominal clitic. As described in section 2.1, the verbal complex acts as a unit. When the verb moves, all the preverbal members of the complex move with it. Similarly, when the subject moves, the subject agreement marker remains in its original position, never moving with the subject. In this regard Swahili SA behaves like Cahuilla SA. Word order is fixed with respect to the verb, suggesting that it is an agreement marker rather than a pronominal clitic.

Summarizing, while a pronominal analysis has been proposed by various authors, the evidence that SA is a pronoun is weak and unclear. The evidence that SA is agreement, on the other hand, is considerably stronger. The arguments for this latter position include the fact quantifier DPs may occur in subject position (showing this position to be subject position and not a topic position), and that the SA prefix behaves like agreement in other languages, as opposed to a pronominal clitic. Thus I conclude that SA is agreement between the subject and the verb.

2.3. *Null Subjects as pro*

Certain languages allow null subjects (e.g., Italian, Spanish) while others do not (e.g., English).

(15)   *He said that [e] is eating   English

(16)   Ha   detto che [e] mangia   Italian
         have-3s said that   eat-3s
         'He said that (he) is eating'

In section 2.1, I showed that Swahili allows null arguments. In this section we will investigate null subjects in Swahili. I will first explain the general theory of null subjects, discussing the licensing condition and the identification requirement on *pro*. I will then show that the null subject in Swahili is *pro*, as in Italian. Swahili null subjects, in addition to satisfying the licensing and identification requirements on *pro*, show other similarities to Italian null subjects.

The fact that null subjects occur in finite clauses in Italian means that the null element is not PRO (as PRO only occurs in non-finite

8

contexts). Furthermore, the null subject in the example above has a definite specific reference, as opposed to an antecedent-controlled reference (as with PRO) or an arbitrary reference (as with PRO$_{arb}$). Null subjects are thus more akin to overt pronominals. Chomsky (1982) concludes that the null element in subject position in a clause such as (16) is the null counterpart to regular pronouns, and is called *pro*. The distribution of overt pronouns and *pro*, however, are different. For example, in Italian *pro* cannot occur as the object of a preposition, but overt pronouns can:

(17)   * Ho parlato con [pro]   Italian
             have-1s spoken with

(18)   Ho   parlato con lui   Italian
            have-1s spoken with him

It was noted that *pro* occurs in languages that have rich subject-verb agreement such as Italian and Spanish, but not in languages without such agreement such as English or French (Taraldsen, 1978). Furthermore, *pro* occurs in non-subject position in languages that have rich object agreement (such as Pashto, (Huang 1989)), or rich indirect object agreement (as in Welsh, (McCloskey & Hale 1984)), as in (19).

(19)   a.   ma [*pro*] wə-xwar-a   Pashto
                   I   PRF-eat-OA$_{fem\text{-}sg}$
                   'I ate (it-fem)'

          b.   * zə [*pro*] xwr - əm   Pashto
                    I   eat − SA$_{1st\text{-}masc}$
                    'I eat (it)'

Both examples in (19) illustrate sentences with a dropped object. (19a) illustrates that in Pashto the object is null in the presence of object agreement on the verb. In (19b), on the other hand, agreement on the verb is with the masculine subject, not with the object, and in this case omission of the object is ungrammatical. Similarly, in Welsh, omission of the indirect object occurs in the presence of agreement between the preposition and the indirect object (example taken from Harbert, 1992):

(20)   Roedd car yn aros amdano [*pro*]   Welsh
          was   car PRT wait for-Msg
          'A car was waiting for (him)'

Rizzi (1986), capitalizing on these restrictions, concludes that the omission of a pronoun involves rich agreement. However, it has also been noted that some languages with rich agreement do not allow *pro*. For example, German does not allow the omission of referential subjects, but does allow the omission of expletive subjects.

(21)   a.   *[e] will zu Hause bleiben   German
                     want at home to-stay
          '(I) want to stay home'

        b.   [e] klar ist, daß er nicht kommen wird German
                   clear is that he not   come   will
          '(It) is clear that he will not come.'

These facts have generally been interpreted as pointing to the existence of two distinct conditions on null subjects: a licensing condition, and an identification condition (Rizzi, 1986). The licensing condition applies to all null pronouns, while the identification requirement only applies to referential/argumental null pronouns. The licensing requirement states that a *pro* must be licensed by its governing head. In Minimalist terms this can be interpreted as *pro* having Case features that must be checked. As for identification, in order for a noun to be referential, it must be specified for person / number features. Therefore, the identification requirement states that a referential pronoun must get ϕ-features through co-indexation with a case-governing head. In Minimalist terms, the pronoun must have its ϕ-features checked by an appropriate head. Presumably there is a relation between rich agreement and the existence of ϕ-features on that head in order to allow identification, though the exact specification of "rich agreement" remains elusive. Therefore in Italian *pro* is identified because the language has rich agreement, while in English this is not the case.

            identification
           |⎯⎯⎯⎯|
(22)   *pro* Parl-o   Italiano   Italian
        speak-SA_{1s} Italian
        '(I) speak Italian'

           ⎯ no identification possible
          |
(23)   * *pro* speak English   English

How does this solve the problem raised by German? Rizzi claims that German satisfies the licensing requirement, but not the

identification requirement. Thus, German licenses non-referential null pronouns, but because it does not satisfy the identification requirement, null referential pronouns are disallowed.[4]

How does Swahili fit into this typology of languages? We saw earlier that Swahili has both rich subject-verb agreement, as well as object-verb agreement. We also saw that Swahili allows null subjects and null objects. Therefore, it appears as if Swahili satisfies the identification requirement for null pronouns. The pronouns that are omitted may be referential arguments (expletives do not occur in Swahili), and so I conclude that Swahili satisfies the licensing condition as well. I conclude that Swahili null subjects are *pro*, akin to null subjects in Italian and Spanish (see Khamisi, 1988 for further evidence that *pro* in Swahili occurs in subject, object and indirect object positions). Furthermore, Swahili null subjects have many of the characteristics of *pro* in Italian that differentiate it from PRO. For example, both Swahili null subjects and Italian *pro* alternate with overt DPs:

(24)   *pro* alternates with overt DPs (unlike PRO)

a.   Juma/*pro*   a – na – zungumz – a ki – zungu   Swahili
         Juma/ *pro* SA_{3s}–PRES–speak–IND 7–English
         'Juma/*pro* speaks English.'

b.   Gianni / *pro* parl–a   Inglese   Italian
         Gianni/ *pro* speak–SA_{3s} English
         'Gianni / *pro* speaks English.'

Furthermore, in both languages null subjects are possible in matrix, finite clauses (unlike PRO, which only occurs in non-finite clauses):

(25)   *pro* occurs in matrix, finite clauses (unlike PRO)

a.   *pro* a – na –zungumz–a   ki–zungu   Swahili
              SA_{3s}– PRES –speak–IND   7–English
         '*he/she* speaks English.'

b.   *pro* parl – a   Inglese   Italian
              speak–SA_{3s} English
         '*he/she* speaks English.'

---

[4] Additionally, there are languages such as Chinese that allow subject and object omission without any agreement whatsoever. This is a problem for Rizzi's proposal. Huang (1984) proposes that these are variables bound by a null topic operator. We shall return to this point in a later section when examining Swahili [-SA] clauses.

Similarly, null subjects in Swahili and *pro* in Italian can both occur in finite embedded clauses (unlike PRO, which occurs in non-finite embedded clauses):

(26)   *Pro* occurs in embedded finite clauses

a. ni–na–fikiri [kwamba *pro* a–na–zungumz–a ki–zungu]   Swahili
SA$_{1s}$–PRES–think   that   SA$_{3s}$–PRES–speak–IND 7–English
'*I* think [that *he/she* speaks English]'

b. Pens – o [che *pro* parl – a   Inglese]   Italian
  think–SA$_{1s}$ that   speak–SA$_{3s}$ English
  '*I* think [that *he/she* speaks English]'

In the next section, we will see that while tense and Mood omission do not occur in Nairobi Swahili, SA omission is in fact permissible. I will propose an analysis that involves a null constant (Rizzi, 1992) in subject position, bound by a topic operator, thereby accounting for null subjects in contexts in which the traditional identifier ('rich agreement') is absent.

3. OMISSION OF VERBAL AFFIXES IN STANDARD SWAHILI

Nairobi Swahili differs from Standard Swahili (Kiswahili Sanifu) in several important ways. I will not discuss all these differences, but will limit myself to discussion of the one difference that is relevant to this paper.[5] In Standard Swahili declarative clauses, the verbal complex is minimally composed of SA-T-V-MOOD. Any omission of any of these elements is considered ungrammatical.   (27)-(29) are the unattested counterparts to (6):[6]

(27)   *   Ø   li - m – pig - a Mariam   SA Omission
          ZERO-SA PAST- OA$_{3s}$- beat - IND Mariam
          'He beat Mariam.'

---

[5] See Deen (2002b) for discussion of this and other differences.
[6] Object agreement differs from subject agreement in that it is not obligatorily present in every declarative utterance. There has been considerable discussion in the literature of the contexts in which OA is obligatory, with some arguing that OA is dependent on animacy (Wald 1979, Hyman & Duranti 1982, Keach 1995), and others arguing that it is dependent on topichood (Hyman & Duranti, 1982) or hearer status (Seidl & Dimitriadis, 1997). In Deen (2002b) I argue that OA is obligatorily present when the object is specific, but obligatorily absent when the object is non-specific.

(28)   * a   -   Ø   - m – pig – a Mariam   T Omission
         SA_{3s} - zero-T - OA_{3s} – beat – IND Mariam
         'He beat Mariam'

(29)   * a - li - m – pig – Ø   Mariam   Mood Omission
         SA_{3s}-PAST-OA_{3s}–beat–zero-MOOD Mariam
         'He beat Mariam'

   In Standard Swahili, the only apparent omission of subject agreement is in the case of the present tense marker –a– (an allomorph of the more common -na-). Such 'omission' occurs only when SA is 3rd person singular:

(30)   a – kul – a   m – kate
         PRES–eat–IND   3–bread
         'He/she is eating bread.'

   I do not consider this a case of omission, but rather phonological reduction because this occurs only when SA and T are homophonous, i.e., in (30) both SA and T are of the form [a]. Thus if SA is of a different person/number specification than 3rd singular (as in 31), such reduction is not possible. Similarly, if the –na– allomorph is used instead of –a– (as in 32), then such reduction is not possible.

(31)   a – kul – a   m – kate
           PRES–eat–IND   3–bread
         * 'He/she is eating bread.'
         * 'I am eating bread.'

(32)   * na – kul – a   m – kate
           PRES–eat–IND   3–bread
         'He/she is eating bread.'

   Thus I assume that this is a phonological reduction of two identical, adjacent segments.[7] Our conclusion therefore is that in Standard Swahili omission of SA or T is ungrammatical.

---

[7] The resulting prefix is usually not a lengthened vowel (Thomas Hinnebush, p.c.), although in the corpus that I use there are examples of lengthened *a*-reduction, i.e., cases in which 3rd person singular SA followed by -*a*- present tense occurs as a single lengthened vowel as opposed to a single short vowel. This variability is likely due to dialectal variation. Further study is required on this issue.

In Nairobi Swahili, however, SA omission occurs occasionally. In a naturalistic corpus of child-directed speech of four Nairobi Swahili speaking families, omission of SA occurs at a significant frequency. I will now discuss this corpus, the omission of SA that is found within this corpus, and then propose a syntactic analysis of such clauses.

4. SA OMISSION IN NAIROBI SWAHILI

The data was collected from January 2000 to November 2000 in Nairobi, Kenya. The participants were four Swahili speaking children and their primary caregivers. The purpose of the data collection was to gather naturalistic data from the four children for analysis of the acquisition and development of their language. Both the adult and child utterances were transcribed in CHAT format, and the child utterances were coded for further analysis using CLAN computer programs (MacWhinney, 2000). Some files were chosen for coding of the caregiver utterances as well, and it is these utterances that form the data for this paper.

During the course of transcription, it was discovered that adults speaking Nairobi Swahili allow the omission of SA. As illustrated in section 3, in Standard Swahili, SA is obligatory in declarative, indicative clauses. Limiting ourselves to this environment, there were a total of 1470 declarative indicative clauses that were coded. Of these, a total of 72 occurred without SA (4.9%). The number of clauses that were missing any other "obligatory" affixes (i.e., tense or mood) were also counted: of the 1470 clauses, 18 were missing an affixes besides SA. The results are presented in Table 4.[8]

Table 4. Omitted affixes in Nairobi Swahili Corpus.

Full Clauses	[-SA] clauses	[-T] clauses	Bare Stems	[-Mood] clauses	Total
1380	72	14	4	0	1470
(93.9%)	(4.9%)	(0.9%)	(0.3%)		

The labels in the above table correspond to the following clause types:

---

[8] In these counts I did not include OA or other grammatical function changing suffixes.

(33)

a.	Full Clause	SA – T – V – IND
b.	[-SA] Clause	Ø – T – V – IND
c.	[-T] Clause	SA – Ø – V – IND
d.	Bare Stem	Ø – Ø – V – IND
e.	[-Mood] Clause	SA – T – V – Ø

Full clauses include all clauses in which SA, T and the Mood final vowel occurred on the verb. OA was not considered, and thus these clauses may or may not include OA. [-SA] clauses are missing SA only, while [-T] clauses are missing the tense marker only. Bare stems are verbs on which neither SA nor T occur. [-Mood] clauses (of which there are none) are clauses that are missing the mood final vowel.[9] Thus SA is the most frequently omitted inflectional affix in Nairobi Swahili.

As mentioned earlier, these data come from Child Directed Speech (CDS), and so an initial hypothesis was that this is peculiar to CDS. However, CDS is generally thought to be a facilitating mechanism for children to acquire the ambient language. In other words, when CDS is different from normal everyday interaction, it is usually simplified, hyper-corrected, fully grammatical speech with exaggerated prosodic contours and facial expressions (see Gleitman, Newport & Gleitman, 1984; Kemler-Nelson *et. al,*1989). It would be unusual and potentially problematic for the children if parents were to produce a model that is blatantly ungrammatical, as the omission of SA is in Swahili. Thus, I conclude that SA omission is a genuine option for adults speaking Nairobi Swahili.

Carol Meyers-Scotton (p.c.) informs me that SA omission is fairly frequent in the spoken forms of many dialects of Swahili. In fact, she documents such a phenomenon (Scotton, 1969) in the Swahili dialects of Baganda and Baluhya speakers in the 1960s. She finds that they frequently omit SA prefixes and rarely use incorrect SA (p. 106). She gives examples such as the following, in which the first line is the dialect, and the second line indicates the Standard Swahili equivalent:

---

[9] In counting [-Mood] clauses, I counted a missing mood vowel irrespective of whether any other prefixes were omitted or not. For example, a [-SA] clause that was also missing a mood final vowel would have been included in the [-Mood] category. As it turns out, the final vowel was never omitted. This is not unexpected since Swahili has phonotactic restrictions that prevent words from ending in a consonant. Thus, a missing final vowel is syntactically as well as phonotactically ill-formed.

(34)   a. mi   na – sem – a   Ø- ta–kuw–a dereva   (dialect)

           mimi ni–na – sem – a   ni–ta–kuw–a dereva (standard)

           I   SA$_{1s}$–PRES–say–IND SA$_{1s}$–FUT–be–IND driver

           'I am saying that I will be a driver.'

     b.   Ø- li – chez – a   m – pila

                ni – li – chez – a   m–pira.

                SA$_{1s}$–PAST–play–IND 3–ball

                'I played ball.'

   Nothing in the way of quantitative data is given, and the context for such omission is not indicated. Meyers-Scotton confirms that there have been no corpus based studies to verify this. Duran (1975) also notes that Kipsigi speakers of Swahili allow SA omission (p.76), as well as various non-standard SA markers. However, no quantitative data are provided here as well, so we do not know how prevalent this phenomenon is. This shows that the phenomenon of SA omission is not restricted to Nairobi Swahili, but also occurs in other dialects of Swahili.

4.1. *Properties of [-SA] clauses*

In the preceding section I showed that [-SA] clauses occur in naturalistic speech in Nairobi Swahili (as well as other regional dialects of Swahili). In this section I will discuss some important syntactic properties of [-SA] clauses. I will show that SA omission is not restricted exclusively to these clauses in Nairobi Swahili, i.e., SA omission does occur in a peripheral set of constructions that have never been brought together in a single description. If I am right in assimilating these various constructions, this suggests that SA omission is a process that is more widely available than previously thought. I will discuss various other constructions that allow (or require) the omission of SA, and focus on two constructions: the habitual and the continuative. I will draw parallels between the continuative and the [-SA] clauses found in the Nairobi Swahili corpus, and conclude that they are the same construction. I will then show that [-SA] clauses allow null subjects: a fact that is unexpected under a theory of identification of null subjects (Rizzi, 1982; Chomsky, 1986). I adopt a proposal of Rizzi's (1992) in which the subject position can be occupied by a null element that Rizzi calls a *null constant*. This null constant is bound by an anaphoric topic operator, which can be either overt or null. I justify this by showing that in [-SA] clauses, the preverbal DP that we see is not a subject but a topic (unlike full

clauses). I then discuss some further characteristics of [-SA] clauses in light of this proposal.

In the Swahili literature, four verbal constructions have been described as allowing/requiring the omission of SA:

- Imperatives
- Infinitives
- Habituals (marked by *hu*)
- Continuatives (marked by *ka*).

I discuss each construction in turn.

4.2. *Imperatives*

(35a) below is an imperative, in which SA is obligatorily absent (Ashton, 1947; Polomé, 1967; Brain, 1960; Myachina, 1981; Schadeberg; 1984):

(35) a. Pig – a picha!   Imperative
          Hit – IND picture
          'Take a picture!'

       b. * **U** – pig – a picha!   * Imperative with SA
           SA₂ₛ – hit – IND picture

Cross-linguistically, imperatives are often (although not necessarily) unmarked for subject agreement (see Koopman, 2001). This appears to be a quite general property of imperatives, an explanation of which goes beyond the scope of this paper. I will thus not discuss SA omission in this context any further, having noted that SA omission is required in imperatives (see Potsdam, 1995; Platzack & Rosengren, 1997; Zanuttini, 1997; Koopman, 2001 for discussions of the absence of agreement in imperatives). I will discuss the other three types of clauses: infinitives, habituals and continuative clauses, and later extend the analysis of continuative clauses to a general phenomenon of SA omission. I will then describe three significant differences between habituals and continuative clauses. I will use these differences to argue that the syntactic structures associated with these two clause types are significantly different.

4.3. *Infinitives*

Neither overt subjects nor SA can occur with infinitives, as examples (36)-(37) show.

(36) a. ni – li – jaribu   ku–end–a   soko – ni   Null subject
       SA$_{1s}$ –PAST–try   INF–go–IND market–LOC   [-SA]
       'I tried to go to the market'

    b. * ni – li – jaribu   **mimi** ku–end–a   soko – ni   Overt Subject
       SA$_{1s}$–PAST–try   **me** INF–go–IND market–LOC   [-SA]
        'I tried to go to the market'

(37) a. * ni – li – jaribu   **ni** – ku–end–a   soko – ni   Null Subject
       SA$_{1s}$ –pres–try   **SA$_{1s}$**– INF –go–IND market– LOC   [+SA]

    b. * ni–li–jaribu   **mimi ni**–ku–end–a soko–ni   Overt Subject
    SA$_{1s}$–pres–try   **me SA$_{1s}$**– INF –go–IND market– LOC   [+SA]

   In these examples, as in their English counterparts, the null subject shares the reference of the matrix subject:

(38)   a. John$_i$ tried [e]$_i$ to go to the market   Subject Control
         b. * John$_i$ tried [e]$_k$ to go to the market

(39)   a. Juma a$_i$–li–jaribu [e]$_i$ ku–end–a soko – ni   Subject Control
            Juma SA$_{3s}$ – PAST –try   INF –go–IND market– LOC
            'Juma tried to go to the market'

         b. *Juma$_i$ a – li – jaribu   [e]$_k$ ku–end–a   soko – ni
            Juma SA$_{3s}$ – PAST –try   INF –go–IND market– LOC

   'Try' is a subject control verb, both in English as well as Swahili, and so I assume that the null element in Swahili is PRO, as it is in English. Additionally, PRO occurs with object control verbs and in arbitrary contexts:

(40)   Mariam$_i$ a – li – mw–omb–a Juma$_k$ [e]$_{*i/k}$ ku–lal–a chini
         Mariam SA$_{3s}$– PAST –OA$_{3s}$–ask–IND Juma   INF–sleep–IND down
         'Mariam$_i$ asked Juma$_k$ PRO$_{*i/k}$ to sleep down (on the floor).'

(41)   Ku – ondok – a mapema si   mzuri
         INF – depart – IND early not good
         'To leave early is not good'

Thus I conclude that PRO occurs in subject position in infinitives in Swahili, as it does cross-linguistically. Thus SA omission in infinitives is neither surprising nor unexplained.

4.4. *Habituals*

Turning now to habituals, consider the examples in (42), taken from (Keach, 1995). We see that subject agreement is obligatorily absent in habitual constructions (cf. 42b, where the presence of SA renders the habitual sentence ungrammatical):

(42)   a. wa – tu wa Kenya hu – wa – pend–a wa – toto
            2-person of Kenya HAB–OA$_2$–like–IND 2–child
            'People of Kenya like children'

         b. * wa–tu wa Kenya   **wa** – hu – wa – pend–a   wa – toto
            2-person of Kenya   **SA$_2$**–HAB–OA$_2$–like–IND   2–child

Furthermore, as Keach (1995) reports, the subject in a habitual clause is obligatorily overt:

(43)   a.   ulevi   hu – ondo – a   akili
                drunkenness HAB–remove–IND sense
                'drunkenness removes common sense'

         b.   * hu – ondo – a   akili
                HAB–remove–IND sense

We will return to an analysis of *hu-* habituals in section 4.6 when we compare habituals and continuatives. We turn now to continuative clauses.

4.5. *Continuative Clauses*

The continuative construction is a regular 'tensed' clause that is used in narratives. It signals an eventuality's continuation in the time line and is marked with the *ka* morpheme. *Ka* occurs in the same position that tense occupies; in complementary distribution with other tense markers.

19

It is thus considered a regular tense in the traditional Swahili literature (see Ashton, 1947; Polomé, 1967)).

(44)   a.   a – ka –kimbi–a   na – o
                  SA$_{3s}$–CONT–run–IND with–rel.
                  '(And then) he ran off with them.'

         b.   * a – li – ka – kimbi – a
                  SA$_{3s}$–PAST– CONT –run–IND

         c.   * a – ka – li – kimbi – a
                  SA$_{3s}$– CONT –PAST–run–IND

   A continuative clause usually takes SA like other tensed clauses, as in (45a) below. However, Ashton (1947) notes that the SA marker may be omitted in certain contexts (cf. 45b, where I have used Ø to indicate that SA has been omitted). She describes the resulting interpretation as expressing 'some emotional quality like mild surprise' (p.134):

(45)   a. a – li – ib – a   wa–toto **a** – ka –kimbi–a   na – o
            SA$_{3s}$–PAST–steal–IND 2–child **SA$_{3s}$**–CONT–run–IND with–rel.
                          'He stole the children and he ran off with them.'

         b. a – li – ib – a   wa–toto Ø ka – kimbi–a   na – o[10]
            SA$_{3s}$– PAST –steal–IND 2–child Ø CONT –run–IND with–rel.
                          'He stole the children and actually ran off with them.'

Thus, continuative clauses may be either [+SA] or [-SA], and may either have an overt or a null subject.

4.6. *Differences between Habituals and Continuatives*

These two constructions differ in several important respects, a few of which we have already seen. In this section I will describe three of these differences: optionality of SA, optionality of subjects, and embedding. I will argue that the omission of SA in habituals results in the lack of an identifier and hence null subjects are prohibited. I will also show that continuative clauses allow null subjects in the absence of SA, which is unexpected given our theory of identification (see section 2.3 earlier). I will then show that continuative clauses are part of a

---

[10] The reference of rel is fixed through discourse. OA is not obligatory in this case because the verb *kimbia* 'run' is intransitive. Substituting a transitive verb in this position such as *piga* 'hit' yields obligatory OA.

broader class of clauses called [-SA] clauses. These clauses are part of colloquial spoken Swahili and are very frequent in child Swahili.

4.6.1. *Optionality of SA*

First, SA in the habitual clause is obligatorily null (as the ungrammaticality of example 44b shows), while the SA in the continuative can be overt (as in example 45a above) or null (as in example 45b above). This null option is pragmatically marked, but in the appropriate contexts, completely grammatical. Judgments on these facts are extremely clear.

4.6.2. *Optionality of subjects*

Second, the subject of the habitual clause must be overt:

(46)　a.　　　ulevi　　　　hu – ondo – a　akili
　　　　　　　drunkenness HAB–remove–IND sense
　　　　　　　'drunkenness removes common sense'

　　　b.　　* hu – ondo – a　　akili
　　　　　　　HAB–remove–IND sense

　　　　　　　　　　　　　　　　(Examples cited in Keach, 1995)

However, the subject of a continuative may be either null or overt. In the unmarked case (when the subject of the continuative clause is the same as the subject of the previous discourse), the subject is null. However, the subject may be overt when there is a change in subject or a clarification is required. For example, in (47a), the subject of the second (continuative) clause is the same as the subject of the main clause. Similarly, in (47b) the subject of the second clause must be the same as the subject of the first clause if the subject is null. When the subject of the second clause is not identical to the subject of the first clause, as in (47c), then an overt subject is required.[11]

(47) a. Juma　　a – li – fik – a　　nyumba–ni　Ø – ka – lal – a.
　　　　Juma SA$_{3s}$–PAST–arrive–IND home–LOC　Ø–CONT–sleep–IND
　　　　'Juma arrived home and (he/*she/*they) then actually went to sleep.'

---

[11] While our eventual goal is to understand SA omission, I am exemplifying here the fact that subjects may be overt or null in continuative clauses in general. This same fact is true in continuative clauses that are missing SA.

21

  b. J na M   wa – li – fik – a   nyumba–ni Ø – ka – lal – a.
   J and M. SA3pl– PAST–arrive–IND home–LOC Ø–CONT-sleep–IND
   'J. and M. arrived home and (they/*he/*she) then actually went to
   sleep.'

c. J. na M.   wa – li – fik – a nyumba–ni. Juma   Ø – ka – lal – a.
   J.and M. SA3pl– PAST–arrive–IND home– LOC J.   Ø–CONT–sleep–IND
   'Juma and Mariam arrived home. Juma then went to sleep.'

So, the subject in continuative clauses may be null or overt depending
on discourse considerations, while the subject in habitual constructions
must be overt.

4.6.3. *Embedding*

A habitual clause can occur in an embedded context as in (48), while
[-SA] continuative clauses cannot, as shown by the contrast in (49a):

(48)   a – li – ni – ambi–a [kwamba wa – tu wa Kenya
      SA3s–PAST–OA1s–tell–IND that   2-person of Kenya

      hu – wa – pend–a   wa – toto]
      HAB–OA2–like–IND 2–child
      'He told me [that people of Kenya like children]'

(49)   a. a –   li – ni – ambi–a [kwamba **a** – ka – kimbi – a]
      SA3s– PAST t–OA1s–tell–IND   that   **SA3s**–CONT–run – IND
      'He told me that he then ran off'

      b. ?? a –   li – ni – ambi–a [kwamba Ø –ka – kimbi – a] [12]
         SA3s– PAST –OA1s–tell–IND   that   Ø –CONT – run – IND
         'He told me that (he) then ran off'

The differences that we have seen so far are summarized in Table 5:

---

[12] My consultant considers this sentence ungrammatical. My judgment is somewhat less
clear, but certainly degraded.

Table 5. Summary of characteristics of habitual and continuative clauses

	SA	Overt Subject	Can be Embedded
Habituals	*	Required	Yes
Continuative	Optional	Optional	No

Recall from the discussion in 2.3 that null subjects must be identified (Rizzi, 1982; Jaeggli & Safir, 1989). Identification can occur through several means, such as control (in the case of PRO), the presence of a c-commanding antecedent (in the case of a trace), or through rich agreement features on a licensing head. This rich agreement provides an identifier for null *pro*. We saw earlier that the null subject of a Swahili tensed clause is *pro*. However, in the case of habitual and [-SA] continuative clauses, this rich agreement is missing. Therefore the question arises as to what the status of the null subject is in such clauses. Specifically, given that the null subject occurs in the absence of SA, how is the identification requirement satisfied?

We saw in the examples in (46) and the summary in Table 5 that habituals simply do not allow null subjects. Therefore, the answer to the question for habituals is clear: because rich agreement is absent, null subjects are blocked. This is consistent with our theory of identification of null *pro.* I therefore assume that subjects in habitual constructions are structurally in subject position and must be overt because of the lack of an identifier.   However, continuative clauses allow the omission of SA, and in those same clauses a null subject is possible. This is not expected under our theory of identification. Additionally, the fact that [-SA] continuative clauses are not possible in embedded contexts is surprising (cf. examples 49), as embedded *pro* clauses are possible in Italian, as well as in Swahili full clauses:

(50)   *pro*   so   che cosa *pro* hai   detto   Italian
          know-1sts what thing   have-2nds said
        '(I) know what (you) said'

(51)   *pro* ni – na – ju – a *pro* u–li–sem–a nini   Swahili
        SA$_{1s}$–PRES–know–IND   SA$_{2s}$–PAST–say–IND what   Full Clause
        '(I) know what (you) said.'

Thus, while *pro* is attested in Swahili full clauses, we have evidence that the null subject in [-SA] continuative clauses is an empty category of a different sort.

Unfortunately, the descriptive evidence available in the literature as to when SA may be omitted is very limited. In order to gain a better empirical understanding of SA omission, I investigated the use of SA

23

by the adults in the Swahili corpus. The first thing I looked at were habitual clauses: not a single utterance containing the *hu-* prefix occurred in the entire corpus. This is presumably related to the fact that the corpus is comprised of child-directed speech, which generally does not contain *hu-* habitual clauses. Secondly, there were no cases of continuative *ka*. This too is most likely due to the context of the recordings. The continuative *ka* is used to tell stories, and the purpose of these recordings was to elicit stories (or any speech) from the children.[13] Thus, continuative *ka* never occurred in the recordings. However, I discovered that SA was omitted in a significant proportion of adult speech in non-continuative contexts. In the next section, I will discuss the contexts of these [-SA] clauses in Nairobi Swahili. I will then provide an analysis of these clauses which postulates a null constant (Rizzi, 1992; 1997) as the null element in subject position. In Deen (2002) I extend this analysis to the speech of children and the omission of SA in child Swahili.

5. [-SA] CLAUSES IN NAIROBI SWAHILI

I conducted a CLAN analysis on the Swahili corpus targeting the adult utterances in 16 files sampled from all four children. I investigated the omission of SA, the expression of tense in these clauses, the identity of the missing referent, and the occurrence of overt subjects in these clauses. Most of the examples that I will provide come from the spontaneous speech of the adult speakers in the Swahili corpus. However, all examples have been verified with a native consultant (as well as my own judgments), and differences in judgments are noted.
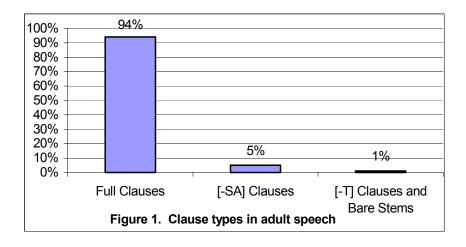
5.1. *Frequency of [-SA] Clauses*

Of the 1470 indicative verbal clauses coded for the adults, 72 (4.9%) are missing SA. Other underspecified clauses (clauses missing tense and clauses missing both tense and SA) account for a combined 1% of indicative clauses. The remaining 94% of indicative clauses are full clauses.

Table 6. Proportions of different clause types in adult Swahili.

Full Clauses	[-SA] clauses	[-T] clauses	Bare Stems	Total
1380 (93.9%)	72 (4.9%)	14 (0.9%)	4 (0.3%)	1470

---

[13] Considering that when recording occurred I explicitly instructed parents that my purpose was to elicit speech from the children. As such, story-telling on the part of the adults was discouraged, as was singing, poetry recitation, non-verbal games, etc.

Figure 1. Clause types in adult speech

A bar chart showing: Full Clauses 94%, [-SA] Clauses 5%, [-T] Clauses and Bare Stems 1%. The y-axis ranges from 0% to 100% in 10% increments.

5.2. *Tense in [-SA] Clauses*

[-SA] clauses occur with a variety of tense markers:[14]

(52)   Ø **na** – tak – a ch–ai?   (Present tense)
        PRES–want–IND 7–tea
        (Hamisi, HAW05)
         '(Do you) want tea?'

(53)   Ø ta – ku – chun – a   (Future tense)
        FUT–OA$_{2s}$ – pinch–IND
        (Mot, MUS10)
        '(I) will pinch you'

(54)   ile ni nini Ø me – lal – a   pa – le ?   (Present Perfect)
        that is what   PR.PERF – sleep – IND LOC – there
        (Joki, HAW01)
        'What is that that has slept over there?'
        (lit: that is what has slept there?)

---

[14] None of the adults used the past tense marker in [-SA] clauses in this corpus. However, my consultant considers the past tense in a [-SA] clause grammatical, and in child speech the past tense marker is used on several occasions.
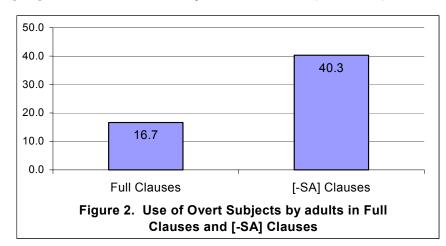
5.3. *Implicit Reference of [-SA] Clauses*

In [-SA] clauses in Swahili, there is no restriction on the implicit reference of the subject. Dropped SA markers can refer to 1st, 2nd and 3rd person referents:

(55) Ø ta – ku – chapa – a   1st sing.
          FUT–OA₂ₛ – slap–IND   (Sam, MUS10)
        '(I) will slap you'

(56) Ø na – ju – a   ku–wach– a   kelele ?   2nd singular
          PRES–know–IND INF–leave–IND noise   (Ala, MUS09)
        '(Do you) know how to stop making noise?'

(57) n–dege   Ø na – ruk – a   3rd singular
        9-bird   PRES–fly up–IND   (Ali, FAU07)
        'The bird is flying up'

   However, there is an asymmetry between [-SA] clauses that have null subjects and [-SA] clauses that have overt subjects. Of the 43 [-SA] clauses that occur with a null subject, the reference of 39 could be determined from context, of which all 39 referred to 1st or 2nd person. Of the clauses that had an overt subject, the subjects were a mix of all three persons. We will return to this point later in the paper.

5.4. *Overt Subjects in [-SA] Clauses*

I conducted a CLAN count of the subjects in full clauses and [-SA] clauses in the Swahili corpus. The results show that in full clauses, adults use overt subjects 16.7% of the time (230 out of 1380) while in [-SA] clauses, adults use overt subjects 40% of the time (29 out of 72).

Figure 2. Use of Overt Subjects by adults in Full Clauses and [-SA] Clauses

Our theory of identification predicts that null subjects should be completely absent in [-SA] clauses because of the absence of an identifier. Therefore the fact that subjects do not increase to nearly 100% is surprising. In fact, null subjects are still the predominant form in [-SA] clauses – a fact that our theory of identification cannot account for. Below are examples of [-SA] clauses with overt subjects as well as with null subjects:

(58) a. wewe Ø ta–kul – a ch–akula? Overt Subject
You FUT–eat–IND 7–food (Ala, MUS08, line 230)
'Will you eat food?'

b. ndege Ø na – ruk – a Overt Subject
bird PRES–climb–IND (Ala, MUS12, line 2372)
'The bird is climbing.'

(59) a. ndio, Ø ta – i – beb – a Null Subject
yes FUT–OA–carry–IND (Ali, FAU01, line 178)
'Yes, (I) will carry it.'

b. Ø na – tak – a ice Null Subject
PRES – want– IND ice (Ham, HAW05, line 135)
'Do (you) want ice?'

In the next section we will investigate how null subjects can occur in [-SA] clauses, given that agreement is generally seen as necessary to identify null *pro*.

6. NULL SUBJECTS IN [-SA] CLAUSES

In the last section we saw that Swahili has a class of clauses in which a null subject appears without an identifier. The primary characteristics of these [-SA] clauses are given in (60):

(60) a. SA omission is optional (4.6.1).
b. Overt subjects alternate with null subjects (4.6.2).
c. They cannot occur in embedded context (4.6.3).

These clauses occur relatively infrequently when compared to full clauses, but when they do occur, they occur primarily with null subjects. We conclude that the null subject is not *pro*, but some other null element that receives identification through some means other than agreement. This null element is syntactically active, as seen in the

27

following [-SA] examples. In (61), the null subject is the antecedent to the reflexive prefix *–ji-*. In (62), the null subject is the controller for the embedded PRO.[15]

(61)   Ø na – ji – on – a
          PRES – REFL. – see – IND
        '(I) see myself.'

(62)   Ø$_i$ na – ju – a PRO$_i$ ku – onge – a?
          PRES – know – IND INF – speak – IND
        'Do (you) know how to speak?'

   The inventory of null elements permitted by UG includes: *pro*, PRO, NP-trace, wh-trace. We have already seen that *pro* cannot be the subject for [-SA] clauses, so we will now consider whether any of the other null elements are possible subjects for [-SA] clauses. By process of elimination, I will show that none of these null elements satisfy the properties in (60). I will then argue that the null element is a null constant (Rizzi, 1992) bound by a topic operator, the details of which will be made clear.

   Let us begin by discussing PRO. PRO is the null element that occurs in the subject position of certain non-finite clauses:

(63)   a.   I entered the race [PRO feeling strong and confident]
        b.   PRO to win the race is important.
        c.   John tried [PRO to win the race]

   We saw earlier that PRO occurs in non-finite clauses in Swahili. However, we can rule PRO out from [-SA] clauses for three reasons. First, PRO occurs prototypically in embedded clauses, and as we saw in (60c), [-SA] clauses do not occur as embedded clauses. Second, PRO occurs in tenseless clauses, while [-SA] clauses always occur with Tense (cf. Examples 52-54). Third, PRO does not usually alternate with overt DPs:

(64)   a. I entered the race [PRO/*Me feeling strong and confident]
        b. PRO/*John to win the race is important
        c. John tried [PRO/*John to win the race]

---

[15] The reflexive example is a constructed example, and the PRO example is an actual utterance from the Swahili corpus (MUS09, line 131). Both are judged grammatical by native consultants.

We saw that in [-SA] clauses, subjects can be overt or null. This is unexpected if the subject is PRO.

Next, let us consider NP-trace. We can eliminate an NP trace from consideration because NP-traces do not alternate with overt DPs either:

(65)   a.   John$_i$ seems [t$_i$ to have left]
          b.   *John seems [he to have left]

Furthermore, NP-traces must be antecedent-bound in order to fulfill the ECP:

(66)   a.   John$_i$, I like [t$_i$]
          b.   * I like [t]

We saw earlier that approximately 60% of [-SA] clauses have a null subject with no overt preverbal DP. Therefore, if the null subject is an NP trace, 60% of [-SA] clauses do not contain an antecedent DP that could bind the NP trace:

(67)   [t] ta –end–a   koti–ni
          _____| fut–go–IND koti–loc
      no antecedent   '(I) will go to court'

This violates the ECP, and should result in ungrammaticality, contrary to fact. Therefore, the null subject cannot be an NP-trace.

A wh-trace has the properties of a variable (Lasnik & Stowell, 1991; Haegemann, 2000). If the null element in subject position in a [-SA] clause is like a wh-trace, it should have the properties of a variable, for example it can be bound by a quantificational element. We see that in [-SA] clauses quantified antecedents are either ungrammatical or marginal at best:[16]

(68)   a.   Kila mw-anafunzi   a – na – som – a   ki – tabu
                 Every 1-student   SA$_{3s}$–PRES–read–IND 7–book
                 'Every student is reading a book.'

---

[16] There is variation in judgments on this point, as my consultant disallows all quantificational antecedents to [-SA] clauses, but I find wh- antecedents marginal and other quantifiers ungrammatical. An additional test for a varable is whether it is sensitive to weak crossover effects (Lasnik & Stowell, 1991). Swahili does not have wh-movement, and so this is difficult to test.

      b.   * Kila mw-anafunzi ∅ na – som – a   ki – tabu
                 Every 1-student   PRES –read–IND 7–book

(69)   a.   Wa–tu   w–ote wa – na – pig – a kelele
                 2-person 2-all SA_{3pl} – PRES – hit – IND noise
                 'Everyone is making noise'

      b.   ??/* Wa – tu w–ote ∅ na – pig – a   kelele
                   2-person 2-all   PRES –hit–IND noise

This suggests that the null element in subject position is NOT a variable, and thus cannot be a wh-trace.

   Summarizing, we have found that the null element in subject position of a [-SA] clause cannot be *pro* (no identifier), it cannot be PRO (doesn't occur in embedded clauses), it cannot be an NP trace (NP traces do not alternate with overt DPs), and it cannot be a wh- trace (it cannot be bound by a quantified antecedent).

<h3 style="text-align:center">7. RIZZI'S NULL CONSTANT</h3>

Rizzi (1992), following Lasnik & Stowell (1991), proposes a new type of null element: a *null constant*. He defines a null constant as:

> - a definite description
> - [–anaphoric, –pronominal]
> - a non-variable
> - an R-expression

   While overt definite descriptions are free to pick up their referent from the discourse, the null version is subject to the identification requirement that all null elements are subject to. He distinguishes the null constant from a null variable. A null variable must be chain connected to a true quantifier for identification, while a null constant (which is –variable) must be chain connected to a non-quantifier (because the Bijection Principle (Koopman & Sportiche 1982; Chomsky 1986) bars vacuous quantification). Thus, a null constant cannot be assimilated to a wh-trace, for example. According to Rizzi, this non-quantifier is typically a null anaphoric operator in an A'-position. The operator cannot be in an A- position because the null constant is an R-expression and thus cannot be A-bound. Thus, the structure proposed by Rizzi is as follows:

(69)   [$_{TopP}$ OP [$_{IP}$ nc [$_{VP}$ …]]]
           |_____|

This binding relation allows identification of the null constant, while also providing a link into the sentence for the anaphoric topic operator.

   This structure holds for modern colloquial German in which it is possible to drop a main clause subject in a V2 construction, i.e., from spec-CP. Rizzi shows that while this is possible in main clauses, the possibility disappears in embedded clauses or when Spec, CP is filled, whether V2 has applied or not (Rizzi's examples):

(70)   a.   *(Ich) habe es gestern gekauft*   German
                '(I) have it yesterday bought'

         b.   *Wann hat *(er) angerufen*?
                'When has he telephoned?'

         c.   *Hans glaubt *(ich) habe es gestern gekauft.*
                'Hans believes I have it yesterday bought.'

         d.   *Hans glaubt daβ *(ich) es gestern gekauft habe.*
                'Hans believes that   I it yesterday bought have.'

Interestingly, the omission of arguments extends to objects as well:

(71)   *(Das) habe ich gestern gekauft.*   German
         'This have I yesterday bought.'

Rizzi notes that this has led researchers to conclude that the examples in (70) and (71) involve topic drop (Ross, 1982), as shown in the structures below.

(72)   a. [$_{CP}$ OP *habe* [$_{IP}$ nc *es gestern gekauft*]]   (=70a)
         b. [$_{CP}$ OP *habe* [$_{IP}$ *ich* nc *gestern gekauft*]]   (=71)

   However, Rizzi notes a fact first pointed out by Cardinaletti (1991), that in colloquial German there is an asymmetry between subject drop and object drop. Cardinaletti claims that 'subject drop can involve pronouns of any specification, while object drop is restricted to 3rd person'. She claims that 3rd person specification is a property inherent to operators, and concludes that subjects should not include an operator. Rizzi therefore limits the above structure (69) to object

omission, and proposes that the structure for German subject omission is as follows:

(73)   [$_{CP}$ nc *habe* [$_{IP}$ t *es gestern gekauft*]]

Thus, the null constant is in the specifier of CP, binding an NP-trace in spec-IP position. Since this structure involves no operator at all, the limitation to 3rd person is removed.

Assuming this structure in (73), Rizzi now must explain how a null element (the null constant) can occur in the structure in violation of the identification requirement.   He claims that the identification requirement is basically the ECP, stated below:

(74)   Empty Categories must be chain-connected to an antecedent.

The structure in (73) violates the ECP as stated above, but Rizzi proposes an addition to the ECP:

(75)   Empty Categories [-P] must be chain-connected to an antecedent… if they can be.

He invokes a notion of the "privilege of the root", whereby elements that ordinarily require binding are exempt from this requirement because they are in the root clause in a position that cannot be c-commanded. Therefore, the null constant in structure (73) is exempt from the identification requirement because it is in the specifier of the root and thus *cannot* be clause-internally identified. He suggests that in this case identification occurs through discourse.

8. NULL CONSTANTS IN SWAHILI

I will adopt Rizzi's proposal for Swahili, and show that [-SA] clauses involve a null constant bound by a topic operator. However, I will diverge from Rizzi's analysis with respect to the question of reference. Instead I will claim that the restriction of object drop to 3rd person referents in German comes through discourse identification restrictions (Gutman, 1999), not syntactic restrictions on the discourse operator. I will adopt the following structure:

(76)

```
              TopP
             /    \
      Operator    Top'
          |         \
          |        AgrSP
          |        /   \
          |      nc    AgrS'
          |_____  /   \
                     Ø    TP
                          /  \
                        T'
                       /  \
                     vP
```

   Assuming a structure as in (76) for Swahili [-SA] clauses, there are various issues that must be resolved. First, we must account for the various characteristics of [-SA] clauses. These characteristics are summarized in (77):

(77)   [-SA] clauses have the following characteristics:
          a.   Cannot occur in embedded context
          b.   Subject can be overt or null
          c.   Can occur with all tenses
          d.   Subject cannot be a quantifier

Additionally, we must resolve the question of what identifies the null constant in the absence of SA or any c-commanding antecedent.

8.1. *Accounting for [-SA] Characteristics*

Let us first consider how this structure can account for the characteristics of [-SA] clauses given in (77a-d), returning in section 8.2 to the question of identification. The first characteristic is that [-SA] clauses never occur in embedded clauses. Under earlier theories of the left-periphery, this result could be derived through the fact that the operator occupies the spec-CP position, and so is in complementary distribution with complementizers. However, under Rizzi's (1997) articulated left periphery hypothesis, this is no longer tenable. Instead, I propose that the restriction to root clauses is due to the nature of the topic operator. The operator is an anaphoric topic operator, and thus looks to discourse for a topic antecedent. If embedded as a syntactic

33

complement, it does not have direct access to discourse, and so cannot occur in such a configuration. It must therefore be in some specifier of the root. Evidence for this comes from the fact that the operator is optionally null[17] – a typical characteristic of the root (Rizzi, 1997).

This leads to the second characteristic: the subject can be either overt or null. When we refer to the 'subject', we are referring to the preverbal DP, which in this case is the anaphoric topic operator. Rizzi's (1997) description of this construction is as follows:

> …the licensing of null constants is not freely available, but is restricted to a designated kind of A'-binder, the anaphoric operator (an element inherently characterized as an operator but different from quantificational operators in that it does not assign a range to its bindee; rather, the anaphoric operator seeks for an antecedent, to which it connects its bindee); anaphoric operators are typically but not necessarily null.
>
> Rizzi, 1997; p.293

Rizzi thus describes a three-member chain (discourse antecedent – anaphoric operator – null constant) in which the anaphoric operator can be optionally null or overt. He describes this as a parametric distinction that some languages allow and others do not, but I propose that Swahili allows both options.

Third, [-SA] clauses can occur with all tenses. This is unsurprising in the structure in (76) as the exact specification of tense is irrelevant to anything in the structure.

The fourth characteristic is that the subject cannot be quantificational. The anaphoric operator is different from a quantificational operator, in that it 'does not assign a range to its bindee'. Therefore the anaphoric operator cannot be quantificational, and since it is the anaphoric operator that surfaces as a preverbal DP, it follows that the subject is not quantificational. It is a property of topics in general that quantification is disallowed (Rizzi, 1997), and so it follows that since the operator is in topic position, quantification should
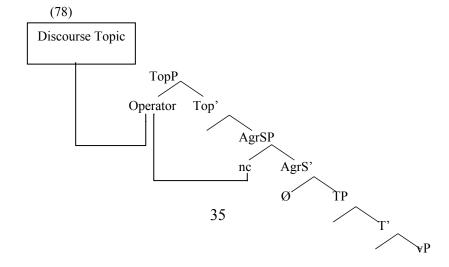
---

[17] The question of *when* the operator can be overt or null is left open at this point. Presumably this turns on discourse conditions, as Rizzi suggests. A clearer understanding of what these conditions are is obviously important, but I must leave it for future study.

not be possible. We have thus accounted for the four characteristics of [-SA] clauses with the proposal that they contain an anaphoric topic operator that licenses a null constant in subject position.

8.2. *Identification in [-SA] Clauses*

Let us now consider the identification requirement.   In the configuration in (76) above, the null constant has no SA to check its Φ-features, and thus is not identified through this kind of feature checking.   The only other possibility is identification through the Operator in spec-TopP position. However, we are now faced with a conflict with Rizzi's claim (from Cardinaletti, 1991) that operators are intrinsically restricted to 3rd person reference.   We saw earlier that reference in [-SA] clauses is not restricted to 3rd person subjects, but is free to refer to all persons. How can we resolve this conflict?

The answer lies in the nature of the operator. As Rizzi himself alludes, it is an anaphoric topic operator. The purpose of an anaphoric topic operator is to provide a link for the discourse topic into the internal structure of the sentence. Therefore, a topic operator links the reference (Φ-features) of the discourse topic to its bindee within the sentence (cf. Huang's 1984 proposal for null arguments in Chinese, a 'discourse-oriented' language). Indeed all topics require this link into the sentence, whether the link is through a trace or other means. In this case, the link is through the binding relation with the null constant. The topic operator receives its Φ-specification from the discourse, and then through a process of feature matching, checks the feature specification on the null constant. This provides identification for the null constant, and it provides a link into the sentence for the discourse topic, via the operator.   Therefore, a more accurate structure of this process is as represented in (78) below:

(78)

```
   ┌───────────────┐
   │ Discourse Topic │
   └───────────────┘
              TopP
              /  \
        Operator   Top'
                   /  \
                 AgrSP
                 /  \
               nc   AgrS'
                    /  \
                   Ø    TP
                        /  \
                       T'
                       /  \
                         vP
```

35

Rizzi (1992) proposes that the discourse operator is intrinsically 3rd person, hence the restriction of object drop in German to 3rd person. In the structure above, the discourse operator has no intrinsic features of its own, but rather gets those features from the discourse topic. Therefore there is no restriction to 3rd person. How do we account for the restriction in colloquial German that Rizzi refers to? While a full explanation of German object drop is not possible here, I believe there is good reason to think that the restriction on object drop is due to discourse constraints, not syntactic ones.

As we will see in the next section, discourse constraints on empty categories play an important role in restrictions on reference. While object drop in German is restricted to 3rd person, null subjects in Hebrew are restricted to 1st and 2nd person. We will see that the Hebrew restriction on null subjects is due to a preference for topics, subjects, agents and conversational partners. It is plausible that the restriction to 3rd person for German null objects is due to discourse preferences for non-topics, non-subjects, non-agents and non-conversational partners.

In the next section I will present facts about Hebrew null subjects and a theory of discourse identification from Arial (1990) and Gutman (1999). We will see that Hebrew has person restrictions on null subjects that are accounted for by discourse principles. This is important because as we see in Swahili [-SA] clauses, the null operator is not restricted to 3rd person. In fact, we will see that the null subject in [-SA] clauses is restricted to 1st and 2nd person, exactly as in the Hebrew case. The overall conclusion that I wish to argue for is that we need not resort to a stipulation about the nature of the anaphoric operator in order to account for person restrictions in German or Swahili.

8.3. *Ariel (1990) and Gutman (1999)*

Ariel (1990) discusses the fact that in Hebrew (in the past and future tenses) null subjects are limited to 1st and 2nd person only. She attributes the Hebrew facts to *discourse* restrictions on when a null subject is possible. She claims that antecedents to null subjects are defined along a scale of accessibility that is determined by various factors. We will restrict our discussion to two of these factors: saliency and unity. Saliency is the relative importance an entity has in the conversation. The more salient an antecedent is in the context, the more accessible it is. Topics (i.e., discourse topics) are very salient and

hence high on the accessibility scale. Similarly 1st and 2nd persons are more salient than 3rd person (because they are conversational partners). Thus examples (79a-b, taken from Gutman, 1999) are grammatical in the absence of any context because they are 1st and 2nd person sentences, respectively. However, (79c) is ill-formed because in the absence of any supporting context, the 3rd person antecedent is not salient enough to identify the null *pro*.

(79)   a. *pro* nixshalti   ba-mivxan be-historia   Hebrew
               failed-1st-SING in-the-test in-history
                '(I) failed the History test.'

      b. *pro*   nixshalta   ba-mixvan be-Historia
             failed-2nd- SING in-the-test   in- history
             '(You) failed the History test.'

      c. * *pro*   nixshal   / nixshela ba-mixvan   be-historia
            failed-3rd-M-SG /F-SG   in-the-test   in-history
          '(He/She) failed the History test.'

   The Saliency Criterion[18] includes several ordered pairs, of which (80) shows the more relevant orderings. Thus, topics are more salient than non-topics, subjects are more salient than non-subjects, and agents are more salient than non-agents.

(80)   Topics > non-Topics
        Subjects > non-subjects
        Agents > non-agents

   The second factor in determining accessibility is Unity. Unity refers to the level of syntactic/semantic cohesion that exists between two sentences, e.g., conjoined sentences are less (syntactically) unified than a matrix and embedded clause, adverbs can increase semantic unity,

---

[18] The Saliency Criterion includes the following ordered pairs, with the element on the left being more salient.

1st and 2nd persons > 3rd person

subject > object > others

split antecedents interpreted as forming a group > split antecedents not interpreted as forming a group

matrix antecedents > embedded antecedents

discourse-topics > non discourse-topics

antecedents in a Focus Presupposition construction > antecedents not in a Focus Presupposition construction

etc. An antecedent that crosses a more unified boundary is more accessible.[19] For example, (81a) is marginal because the antecedent-*pro* relation crosses a sentence boundary that is not semantically unifed. In (81b), with the addition of semantic adverbials, unity is increased and thus accessibility is increased.

(81)   a. ?? $Noga_i$ rak   hitxatna im $Shimon_j$ Hebrew
              Noga only got-married with Shimon

        ve-kvar   *pro_{i+j}*   hitgrashu
        and-already   got-divorced-pl.

       'Noga just married Shimon, and (they) already got divorced.'

     b. rak   lifney xodesh $Noga_i$ hitxatna   im $Shimon_j$
       only before month Noga got-married-f with Shimon

       we-kvar   ba-shavua   she-avar *pro_{i+j}* hitgarshu.
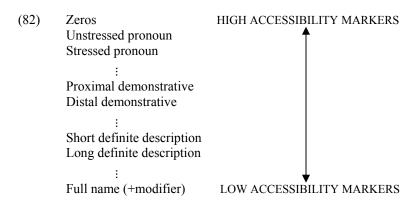   and-already in-the-week that-passed   got-divorced-pl

       'Only a month ago Noga married Shimon, and last week (they) already got divorced.'

  Ariel also assumes that noun phrases differ in the degree to which they depend on antecedents. For example long definite descriptions occur lower on her scale of accessibility than short definite descriptions, which in turn occur lower on the scale than stressed pronouns, etc. At the highest end of the scale are gaps, i.e., null subjects and objects. This is shown in (82). Noun phrases at the higher end of the scale will only recover antecedents with a higher level of accessibility (e.g., topics). Similarly, noun phrases at the lower end of the scale can recover antecedents that are lower in accessibility. Therefore, gaps, which are the highest in the scale, require the highest degree of accessibility, and hence are the most restricted.

---

[19] The Unity Criterion includes the following ordered pairs, with the element on the left being the more unified.
embedding > conjoining
sentences with parallel time-adverbials > sentences with no parallel time adverbials
sentences with consequence adverbials > sentences with no consequence adverbials
sentences with other adverbials > sentences with no other adverbials

(82)   Zeros   HIGH ACCESSIBILITY MARKERS
         Unstressed pronoun
         Stressed pronoun
                    ⋮
         Proximal demonstrative
         Distal demonstrative
                    ⋮
         Short definite description
         Long definite description
                    ⋮
         Full name (+modifier)   LOW ACCESSIBILITY MARKERS

   In the case of null subjects, Ariel finds that topics are more
accessible than non-topics. This means that null subjects are more
likely to occur in contexts in which there is a clear discourse topic
available to the listener as an antecedent. In cases where there is no
topic, a null subject is not discourse identified. Similarly, subjects are
more accessible than non-subjects, and agents are more accessible than
non-agents. Therefore, null subjects seek out topics, subjects and
agents more than other elements tend to act as antecedents for discourse
identification.

   How does this theory account for the Hebrew pattern of subject
drop? In past and future tenses, Hebrew allows subject drop of 1st and
2nd person pronouns, but not 3rd person pronouns (although see below).
Recall that according to Ariel, null subjects are the highest accessibility
markers and thus require an antecedent that is high in accessibility.
Ariel claims that 1st and 2nd person antecedents are inherently more
salient in the discourse than 3rd person antecedents because they are
conversational partners. Because null subjects require the highest level
of accessibility, this reduces the frequency of 3rd person null subjects.
In fact, Gutman (1999) shows that 3rd person null subjects are not
completely unattested, but are considerably less frequent than 1st or 2nd
person null subjects.   Gutman shows that when a sufficient level of
accessibility is created (through increased saliency and unity), Hebrew
allows null 3rd person subjects, as shown in (83).

(83)   Joan$_i$ soxaxa   ita   axshav be-ivrit,   af   ki
         Joan chatted-f with-her now   in-Hebrew, even though

39

*pro*ᵢ hevina   rak xelek min   ha-dvarim she-ha-yalda amra.
understood-f only part from the-things that-the-girl said-f

   'Joanᵢ was chatting with her in Hebrew now, even though
   (sheᵢ) understood only part of what the girl said.'

  Example (83) is taken from a novel, and demonstrates that in an adjunct clause (high in unity) with a matrix subject antecedent (high in salience), pro-drop is possible in the 3rd person. Gutman argues that this is because the antecedent accessibility is extremely high, as well as the fact that this occurs in literate Hebrew. She argues that the literate medium increases macro (or global) accessibility, making pro-drop much easier.[20]

  In addition to this person restriction in past/future tenses, Hebrew disallows null subjects entirely in the present tense. Gutman (1999) argues this is because of an additional condition that impacts null subjects: null subjects must be syntactically identified in order for discourse identification to be possible. Hebrew present tense is unmarked for person features, and so null subjects are not *syntactically* identified. This renders discourse identification irrelevant. In the past/future cases, however, both syntactic as well as discourse identification affect the occurrence of null subjects. Gutman provides a series of sentences in the present tense with increasing levels of accessibility, and we see that in each case, null subjects are disallowed. In (84), the present tense embedded clause does not allow *pro*, despite a subject antecedent.   In (85), the accessibility is increased by incorporating (85) into a conjoined-clause structure with parallel-time adverbials.   Because of the parallel time adverbials, an additional clause must be added.

(84)   * Rinaᵢ hodi'a   *pro*ᵢ she-magi'a   be-shesh
       Rina informed-f   that-arrive-f-sg at-six
       'Rina informed that (she) is arriving at six.'
       (lit: Rina informed that (she) is arriving at six.')

---

[20] She argues that literate contexts inherently increase saliency, since Ariel's original definition of salience was based on processing capacity. Arial argued that the less salient the antecedent the more taxing it is to link to a null argument. In literature, recovery of identity is considerably easier because of the written medium. In fact, even in English in certain literate contexts of extremely high salience, null subjects are grammatical and very usual. For example, "contains 100% fruit juice" found on a product label. Gutman argues that such examples are licit in English because of the extremely high salience of the antecedent – the product on which the label occurs.

(85)   */?? etmol   Rina$_i$ hodi'a   she-hi   magi'a   be-sheva,
           yesterday Rina   informed-f that-she arrive-f-sg at seven

       ve-hayom hi$_i$ hodi'a   *pro$_i$* she-magi'a   be-shesh
       and-today she informed-f   that-arrive-f-sg at six

       'Yesterday, Rina informed that she wouyld arrive at seven,
       and today, she informed that (she) would be arriving at six.'

In (86), Gutman adds a preceding sentence that makes the subject into a topic, thus raising the salience even further.

(86) */? Rina$_i$ hi kol-kax lo   hexletit!
         Rina she   so NEG decisive

    etmol   hi$_i$ hodi'a   she-hi$_i$ magi'a
    yesterday she informed-f that-she arrive-f-sg

    be-sheva, ve-hayom hi$_i$ hodi'a   *pro$_i$* she-magi'a   be-shesh
    at-seven and-today she informed-f   that-arrive-f-sg at six

    'Rina is incapable of making a decision! Yesterday, she informed
    that she would arrive at seven, and today, she informed that (she)
    would be arriving at six.'

    Example (86) has the highest possible level of salience and unity, and still null subjects are disallowed in the present tense. Thus Gutman concludes that while discourse identification is important in Hebrew, syntactic identification must also be satisfied.

    The case of Hebrew highlights the fact that restrictions on person (or number or gender) in null subjects need not necessarily be a result of a syntactic process. In this case we saw that the restriction in Hebrew of null subjects occurring in 1st and 2nd person is due to the saliency of 1st and 2nd person as conversational partners, coupled with the fact that null subjects require a very high level of accessibility.

    Let us now consider Swahili [-SA] clauses. Since [-SA] clauses are null subjects, we expect a similar pattern as we see in Hebrew: null subjects have a tendency towards taking 1st and 2nd person antecedents because they are more salient than 3rd person antecedents. In fact, this appears to be the case in Swahili.

Recall that in section 5.3 we saw that the reference of [-SA] clauses is free. However, in that section we only discussed the reference of overt subjects. Of the 72 [-SA] clauses in the Swahili corpus, 43 occur with a null subject, and 29 occur with an overt subject. Of the 43 null subject [-SA] clauses, the reference of the null subject was determined on the basis of context. 4 utterances were discarded due to unclear reference. Of the remaining 39 null subject [-SA] clauses, all 39 were either 1st or 2nd person. The overt subject [-SA] clauses, as we saw in section 5.3, refer to all three persons. This is shown in Table 7 below (for examples, refer to section 5.3).

Table 7.   Person reference in [-SA] clauses with Overt and Null subjects

	Overt Subject	Null Subject
1st person	6	17
2nd person	7	22
3rd person	16	0

Unclear = 4

We see that restrictions on person occur when the subject is null, but not when the subject is overt. Whether the subject is silent or overt is important in determining any person restrictions in [-SA] clauses. In the absence of an overt operator, the null constant seeks a salient antecedent (in Ariel's terms), hence a preference for 1st or 2nd person (conversational partners). However, if the operator is overt, there are no inherent restrictions on person. Thus, the restriction to 1st and 2nd person for null subjects in [-SA] clauses is due to discourse principles that apply only to null elements, as described by Gutman (1999) and Ariel (1990) for Hebrew, not an inherent restriction on the operator.[21]

Our conclusion therefore is that [-SA] clauses involve a topic operator – null constant construction. The topic operator can be overt or null, but when null we see the effect of discourse principles on the reference of the null subject.

9. CONCLUSION

In this paper we discussed the nature of SA in Nairobi Swahili, concluding that it is agreement between the subject and verb. We then

---

[21] Intuitions about 3rd person null subject [-SA] clauses vary from speaker to speaker. Of the three native Nairobi Swahili speakers that I have consulted (myself included), the consensus is that 3rd person is "confusing", i.e., there is a strong preference to interpret a [-SA] clauses as 1st or 2nd, and forcing a 3rd person interpretation through rich context conflicts with this preference. This is entirely congruous with the proposal in the text that discourse restrictions prevent 3rd person null subject [-SA] clauses.

established that Swahili is a null subject language and that in the case of full clauses, the null subject is *pro*, as in Italian and other null-subject languages. We then discussed clauses that permit the omission of SA. We saw that there are two major types of clauses in Swahili that allow SA to be omitted - habitual clauses that do not allow null subjects, and [-SA] clauses that do allow null subjects. The latter raise a problem for the identification requirement on null elements. I proposed that in [-SA] clauses, the subject position contains a null constant licensed by an anaphoric topic operator. This anaphoric operator seeks out a discourse antecedent, to which it links its bindee (the null constant). It is in this way that the null constant is identified. Its reference is not restricted if the operator is overt, but when null, there is a preference for more salient antecedents.

Like other languages, Swahili has null subjects: *pro*, PRO and null constants. Each of these empty categories has distinct properties and is governed by distinct syntactic as well as discourse rules. In Deen (2002), I look at the development of these different types of null elements in child language, finding that children know the properties of null elements at a surprisingly early age. The exact discourse restrictions on when [-SA] clauses may occur are not well understood at this point, and require further investigation.

REFERENCES

ARIEL, MIRA. 1990. *Accessing NP antecedents*. London, Routledge.

ASHTON, ETHEL OSTELI. 1947. *Swahili Grammar*. London, Longmans, Green and Co.

BRAIN, JAMES. 1960. *Basic structure of Swahili*. Eastern African Studies Program, Syracuse University.

CARDINALETTI, ANNA. 1997. Subjects and Clause structure. In *The new comparative syntax*. L. Haegeman (Ed.) London, Longman**: 33-63.

CHOMSKY, NOAM. 1982. A Note on the Creative Aspect of Language Use. *The Philosophical Review*, 1982, 91, 3478, July, 423-434

CHOMSKY, NOAM. 1986. *Barriers*. Cambridge, MA, MIT Press.

DEEN, KAMIL UD. 2002. *The Acquisition of Nairobi Swahili: The Morphosyntax of Inflectional Prefixes and Subjects*. Unpublished Doctoral Dissertation, UCLA.

DURAN, JAMES. 1975. *The Role of Swahili in a Multilingual Rural Community in Kenya*. Unpublished Doctoral Dissertation, Stanford University.

GLEITMAN, LILA; NEWPORT, ELISSA & GLEITMAN, HENRY. 1984. The current status of the Motherese Hypothesis. *Journal of Child Language*, 11, 1, 43-79.

GUTMAN, EYNAT 1999. *Null Subjects: A theory of syntactic and discourse-identification*. Unpublished Doctoral Dissertation, University of Delaware.

HAEGEMANN, LILIANE. 2000. Inversion, Non-Adjacent Inversion and Adjuncts in CP. *Transactions of the Philological Society*, 2000, 98, 1, May, 121-160.

HUANG, JAMES C.-T. 1984. On the distribution and reference of empty pronouns. *Linguistic Inquiry* **15**: 531-574.

HUANG, JAMES C.T. 1989. Pro-drop in Chinese: A Generalized control theory. *The Null Subject Parameter*. O. Jaeggli and K. Safir Eds. Dordrecht, Kluwer Academic Publishers**:** 185-214.

JAEGGLI, OSVALDO AND KENNETH SAFIR. 1989. *The Null Subject Parameter*. Dordrecht, Kluwer.

KEACH, CAMILLIA. 1995. Subject and Object markers as agreement and pronoun incorporation in Swahili. *Theoretical approaches to African Linguistics*. A. Akinlabi Ed. Trenton NJ, African World Press**:** 109-116.

KEMLER NELSON, DEBORAH; HIRSH-PASEK, KATHY; JUSCZYK, PETER; CASSIDY, KIMBERLY WRIGHT. 1989. How the Prosodic Cues in Motherese Might Assist Language Learning. *Journal of Child Language*, 1989, 16, 1, Feb, 55-68

KHAMISI, A. M. 1988. A Typology of Gaps in Kiswahili sentences. *Kiswahili: Journal of the Institute of Kiswahili Research* **55**1: 120-133.

KOOPMAN, HILDA AND DOMINIQUE SPORTICHE. 1991. The position of subjects. *Lingua* **85**: 211-258.

KOOPMAN, HILDA. 2001. Topics in Imperatives. *Imperatives in Generative Grammar*. W. van der Wurff Ed. Benjamins Publications

KOOPMAN, HILDA AND ANNA SZABOLCI. 2000. *Verbal Complexes*. Cambridge, MA, MIT Press.

KRIFKA, MANFRED. 1995. Swahili. In *Syntax: Ein Internationales Handbuch Zeitgenossischer Forschung. An International Handbook Of Contemporary Research*. J. Jacobs, A. von Stechow, W. Sternefeld and T. Vennemann Eds. Berlin, Walter de Gruyter. **2:** 1397-1418.

LANGACKER, RONALD. 1977. The Syntax of Propositions in Uto-Aztecan. *International Journal of American Linguistics*, 1977, 43, 1, Jan, 11-26

LASNIK, HOWARD & TIM STOWELL. 1991. Weakest Crossover.

Linguistic Inquiry **22**: 687-720.

MACWHINNEY, BRIAN. 2000. The CHILDES project: Tools for analyzing talk. Third Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

MCCLOSKEY, JAMES & HALE, KENNETH. 1984. On the Syntax of Person-Number Inflection in Modern Irish. *Natural Language & Linguistic Theory*, 1983, 1, 4, 487-533

MYACHINA, EKATERINA NIKOLAEVNA. 1981. *The Swahili language: a descriptive grammar*. London, Routlege & Kegan Paul.

PLATZACK, CHRISTER & ROSENGREN, INGER. 1997. On the Subject of Imperatives: A Minimalist Account of the Imperative Clause. *Journal of Comparative Germanic Linguistics*, 1997-1998, 1, 3, 177-224

POLOMÉ, EDGAR. 1967; *Swahili Language Handbook*. Center for Applied Linguistics, Washington, DC.

POTSDAM, ERIC. 1995. Phrase Structure of the English Imperative. *Indiana University Linguistics Club Publications*, 1995, 1, 143-154.

RIZZI, LUIGI. 1982. *Issues in Italian Syntax*. Dordrecht, Foris.

RIZZI, LUIGI. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* **17**: 501-557.

RIZZI, LUIGI. 1992. Early null subjects and root null subjects. *Language acquisition studies in generative grammar*. T. Hoekstra and B. Schwarz Eds. Amsterdam/Philadelphia, John Benjamins**:** 151-176.

RIZZI, LUIGI. 1994. Some notes on Linguistic theory and language development: the case of root infinitives. *Language Acquisition* **3**: 371-393.

RIZZI, LUIGI. 1997. The fine structure of the left periphery. *Elements of grammar. A Handbook of generative syntax*. L. Haegemann Ed. Dordrecht, Kluwer.

SCHACHTER, PAUL. 1995. Tagalog. Chapter in *Syntax: Ein Internationales Handbuch Zeitgenossischer Forschung. An International Handbook Of Contemporary Research, 2. Halbband/Vol. 2,* Jacobs, Joachim, von Stechow, Arnim, Sternefeld, Wolfgang, & Vennemann, Theo [Eds], Berlin, Federal Republic of Germany: Walter de Gruyter, 1995, pp 1418-1430

SCHADEBERG, THILO. 1984. *A Sketch of Swahili Morphology*. Foris Publications.

SEIDL, AMANDA & DIMITRIADIS, ALEXIS. 1997. The Discourse Function of Object Marking in Swahili. *Papers from the Regional Meetings, Chicago Linguistic Society,* 1997, 331, 373-387

STEELE, SUSAN. 1995. Clisis. Chapter in *Syntax: Ein Internationales Handbuch Zeitgenossischer Forschung. An International*

Handbook Of Contemporary Research, 2. Halbband/Vol. 2, Jacobs, Joachim, von Stechow, Arnim, Sternefeld, Wolfgang, & Vennemann, Theo [Eds], Berlin, Federal Republic of Germany: Walter de Gruyter, 1995, pp 1226-1235

TARALDSEN, KNUT TARALD. 1978. *On the NIC, vacuous application and the that-trace filter.* Bloomington.

VITALI, ANTHONY. 1981. *Swahili Syntax.* Dordrecht, Foris Publications.

ZANUTTINI, RAFFAELLA. 1997. *A Comparative Study of Romance Languages.* New York, Oxford University Press.

ZWICKY, ARNOLD AND GEOFFREY PULLUM. 1983. Clticization versus Inflection: English n't. *Language* 59, 502-513