Subject Agreement in Nairobi Swahili

Kamil Ud Deen
University of Hawaii at Manoa

1. Introduction

A current debate in Bantu linguistics centers on whether the subject marking prefix (henceforth SM) is an agreement marker or a pronominal clitic. This issue was brought to the forefront of the field by Bresnan & Mchombo (1987, henceforth B&M). B&M investigated the status of agreement prefixes in Chichewa showing that SM in Chichewa functions both as a pronoun as well as agreement, while Object Marking (henceforth OM) is unambiguously pronominal. Keach (1995) applies their approach to Swahili, concluding that in Swahili (as in Chichewa), SM is ambiguous between being a pronoun and an agreement marker. Related to this, Zwart (1997) argues that the tense marker in Swahili is not a tense marker in the traditional sense, but an auxiliary verb. He argues that SM is a subject pronoun cliticized to this auxiliary verb. The lexical subject, according to Zwart, is in fact not the subject but a topic (see (1) below). Thus Zwart argues for a purely pronominal analysis of SM in Swahili.

Subject / topic? Agreement marker / subject clitic?

(1) Juma a na kimbi a

‘Juma is running.’

In this paper I argue that within a particular dialect of Swahili spoken in and around Nairobi, the capital of Kenya, SM is an agreement marker and not pronominal. In section 1.1, I first explicate the details of ‘an agreement analysis of SM’ and ‘a pronominal analysis of SM’. In section 2 I then describe Keach’s analysis of SM that supports the agreement view in Nairobi Swahili. In section 3 I provide several additional arguments against a pronominal analysis of SM, as argued for by Zwart (1997). I conclude in section 4 with a discussion of how to account for some of the disparities in the literature regarding the status of agreement in Swahili.

1.1 Agreement versus Pronoun

Let us begin by clarifying the difference between an ‘agreement analysis’ and a ‘pronominal analysis’. Agreement is a process in which two elements that are in a local configuration (2a) share morphological features through a process of feature matching (or checking). Although not crucial to

---

*Thanks to Mzee Mchombo for discussion and comments, as well as Leston Buell, Dominique Sportiche, Hilda Koopman, Thomas Hinnebusch, and an anonymous reviewer for useful comments. All errors are my own.

1 OM appears to function differently from SM in both Chichewa as well as Swahili. In Swahili OM is not obligatory, as SM is (although see Deen 2002, 2003a, 2003b for details of contexts in which SM does not occur), and it occurs only when the direct object is specific. There is an intuition amongst native speakers of Swahili that OM is somehow referential, or indexical in nature. This squares well with the results of the functional tests performed by Keach which show that OM is more pronominal than SM. I do not make any claim about OM in this paper, but limit myself to the analysis of SM.

2 Recent work in Minimalism (Chomsky 1998, 2001) has raised the possibility that agreement relations need not be strictly local. I will not consider the possibility of a slightly ‘looser’ agreement configuration because all agreement relations are strictly local in Swahili. In other languages in which non-local relationships occur, we might need to reconsider our understanding of agreement.
my analysis, I assume that the subject raises from a lower VP position to [spec, AgrSP] (Koopman & Sportiche, 1991) and triggers agreement with the head AgrS. This subject can be optionally null, in which case it is identified through rich agreement (Rizzi, 1982; Taraldsen, 1978).

Thus under an agreement analysis (2a), SM is the head of AgrS and the subject with which it agrees is in the specifier. Under a pronominal analysis (2b), SM is a pronominal DP raised from [spec, VP] to the spec of AgrSP. The head of AgrS is phonetically null. Crucially, under a pronominal analysis, in a sentence with an overt ‘subject’ as in (3) below, because subject position is occupied by SM, what looks like the lexical subject is in fact a topic that binds the pronominal clitic in subject position, as shown in (3c).

(3) a. Kibaki a – li – shind – a
   Kibaki SA₃s – past – win – IND
   ‘Kibaki won.’

b. Agreement Analysis:  
   Subject a – li – shind – a
   Subject Agreement – tense – verb – mood

c. Pronominal Analysis:  
   Topic a – li – shind – a
   Topic Subject – tense – verb – mood

Note, therefore, that under a pronominal analysis a preverbal ‘subject’ in Swahili should have the properties of a topic and not a subject (as pointed out by Zwart, 1997). Distinguishing between these two analyses is not an easy task, as a survey of the Bantu literature reveals.

B&M argue that in Chicheŵa, an East Central African Bantu language, SM is ambiguous between agreement and a pronoun. They use several functional diagnostics, including the interplay of word order with agreement morphology and tone with phrase structure. They show that overt subject NPs differ from overt object NPs in their distribution in relative clauses and interrogatives, these being contexts which allow and disallow topics respectively (see the next section for more detailed reasons for why this is so). Chicheŵa differs from Swahili in some significant ways,³ and so the details of their

³ For example, Chicheŵa has tone, Swahili does not. Chicheŵa does not exhibit the ‘Definiteness Effect’ in Object Marking (B&M, 1987, p.761), while Swahili does. Chicheŵa allows postverbal subjects to be questioned, a fact that B&M take to mean that ‘the subject and topic NPs appear at the same level of structure in the S, with exactly the same ordering possibilities’ (p.775). Swahili, on the other hand, disallows such questions:

a. Nani alipiga picha?  b. ?!* Alipiga picha nani?
   ‘Who took a picture?’        ‘Took a picture, who?’
discussion are not directly relevant to us. However, their methodology for distinguishing agreement from pronouns has been applied to Swahili by Keach (1995).

2. Arguments for Agreement

Keach (1995) argues that SM in Swahili, as in Chichewa, behaves as both agreement and a pronoun. Following B&M, she defines agreement as being a local phenomenon (i.e., it always occurs in a spec-head configuration). The process of agreement involves the sharing of Φ-features between the XP (in spec position) and the agreeing marker (the head). A pronominal analysis, on the other hand, is one in which SM is the subject, anaphorically bound by the topic DP. Theta role assignment occurs directly to the SM pronoun, and the theta role is transmitted through a chain to the overt topic DP. Keach presents three data arguments, which yield conflicting results (the first in favor of a pronominal analysis, the other two in favor of an agreement analysis), hence the claim that SM is ambiguous between agreement and a pronoun. I will now present the three arguments put forward by Keach (1995). To anticipate, I reject her argument in favor of a pronominal analysis, but accept her two arguments in favor of an agreement analysis. This therefore resolves the conflicting results, leading us to conclude that SM in Swahili is agreement.

2.1 Habitual Hu-

Keach starts by showing that post-verbal subjects are possible ordinarily, as are subjects raised to the topic position of a higher clause, as in (4b-c, Keach’s 4b-c).

(4) a. watu wa Kenya_[i], wa_[i]– na – wa – pend – a watoto
   people of Kenya SA3pl-pres-OA3pl-like-IND children
   ‘People of Kenya like children’

   b. wa_[i]– na – wa – pend – a watoto watu wa Kenya_[i]
   SA3pl-pres-OA3pl-like-IND children people of Kenya
   ‘(They) like children, people of Kenya’

   c. watu wa Kenya_[i]; ni – na – fikir - i kuwa wa_[i]– na – wa–pend–a watoto
   people of Kenya SA1s-pres-think-IND that SA3pl-pres-OA3pl-like-IND children
   ‘People of Kenya, I think that, (they) like children.’

According to Keach, these examples are compatible with both a pronominal and an agreement analysis. Under the agreement analysis, agreement occurs before movement, and then the subject DP is moved leftward or rightward as normal. Under a pronominal analysis theta role assignment occurs directly to SM, and is then transmitted through a chain to the overt DP. Keach then presents data showing that SM is ungrammatical when the HU- tense marker (indicating habituality) is used:

(5) a. Watu wa Kenya hu - wa – pend – a watoto
   people of Kenya hab-OA3pl like - IND children
   ‘People of Kenya like children’

   b. *Watu wa Kenya wa - hu - wa – pend – a watoto
   people of Kenya SA3pl-hab-OA3pl like - IND children
   ‘People of Kenya like children’

In sentences such as (5a), where SM is absent, postverbal subjects and the raising of subject to matrix topic position are ungrammatical as illustrated in examples (6a,b):
According to Keach, this is evidence that SM is a pronoun for the following reason: theta role assignment occurs directly to the SM pronoun and is then transmitted through a chain to the topic. Eliminating the SM thus results in a theta-criterion violation. No violation occurs when the topic is local and can receive its theta role directly (example 5a). For Keach, a DP in topic position is sufficiently local for theta assignment to occur, but once that DP is moved (either postposed or raised to a higher clause), that local relationship no longer exists. According to Keach, it is unclear how to account for this ungrammaticality under an agreement analysis.

Let us examine the logic of this argument. As discussed earlier, according to Keach, in the case of non-habitual clauses in which the subject has been topicalized into a higher matrix clause or postposed to the end of the sentence (as in examples 4a-c), the theta role is assigned to SM and is then transmitted through a chain from the subject pronoun to the topic DP. This accounts for why postverbal subjects and topicalizing to a matrix sentence are possible. However, it is not clear why in hu-habitual clauses in which the subject has been topicalized into a higher matrix clause or postposed to the end of the sentence (as in examples 6a-b), a theta role cannot be assigned to the trace of the topic and transmitted by the chain that connects it to the moved topic. Both constructions involve the transmission of the theta role via a chain. The difference cannot be due to the different A/A’ status of these chains, as both are A’-chains. Thus the ungrammaticality of postposed subjects in hu-habitual clauses cannot be due to the status of SM as a pronoun, but must be attributed to other factors.

The possibility of postverbal subjects is related to the existence of rich agreement. For example, Italian and Spanish are two languages with rich subject verb agreement and both allow postverbal subjects (modulo certain lexical effects). On the other hand, English does not have rich agreement and does not allow postverbal subjects. We see this correlation within the same language here: in the presence of SM, postverbal subjects are possible, but not in the absence of SM. An analysis of postverbal subjects goes beyond the scope of this paper, but the relevance here is that the ungrammaticality of postverbal subjects in habituals in Swahili does not necessarily entail that SM is pronominal, and given the cross-linguistic correlation of postverbal subjects with agreement, it in fact implicates the presence of agreement.

2.2 Subject Wh- Questions

Keach’s second argument is based on three principles proposed by B&M in their analysis of Chichewa. These principles are stated in (7) below. The first of these principles is not relevant for our purposes. Based on the latter two principles, Keach examines subject wh- questions in Swahili, exemplified in (8).

---

4 Keach provides examples which show that subjects in Hu-clauses must be overt:

i. ulevi hu - ondo - a akili
   drunkenness HAB-remove-IND sense
   ‘Drunkenness removes common sense’

ii. *hu - ondo - a akili
    HAB-remove-IND sense

This suggests that the theta criterion must be satisfied through the subject directly in such cases of missing SM. Thomas Hinnebusch (p.c.) informs me that this may not be as clear as Keach suggests. He reports that native speakers do use hu-clauses in conversational speech without overt subjects. Leston Buell (p.c.) also informs me of finding tokens of this sort in written plays. It is unclear whether this is a reflection of competence or performance, and because at least one native consultant agrees with these judgments, I continue to assume Keach’s data.
(7) (i) Relative pronouns bear TOPIC function;
    (ii) Questioned constituents bear FOCUS function;
    (iii) An argument cannot bear both TOP and FOC function in the same clause.

(8) nani, a – me – end–a ?
who SA
3s-pr.prf-go-IND
who has gone?

Swahili does not have wh- movement, and so the wh-word in (8) is in situ. Under a pronominal analysis, the SM is in subject position and the wh- phrase is in topic position. The wh- phrase therefore bears TOPIC function. Furthermore, because it is the questioned constituent, by principle (7ii), it bears the FOCUS function as well. Principle (7iii) rules this ungrammatical, as the wh- word bears both TOPIC and FOCUS functions. However, as (8) shows, subject wh-questions are possible in Swahili. Keach concludes that, assuming the principles in (7), SM cannot be pronominal.

2.3 Idioms

Keach’s final argument comes from idioms. She notes that idiom subjects resist topicalization, as was pointed out by B&M. In (9b), the idiom subject mtindi ‘brew’ is topicalized to a higher clause, and this results in ungrammaticality, suggesting the idiom subject is a true subject and not a topic. This again argues in favor of an Agreement analysis.

(9) a. Ni – li – fikir - i    kuwa mtindi u - me – va - a Asha
SA1s-past-think-IND that brew SA3-pr.prf.-wear-IND Asha
‘I thought that Asha is drunk’
(lit: I thought that the brew has covered Asha)

Brew SA1s-past-think-IND that SA3-pr.prf.-wear-IND Asha
(lit: (As for) the brew, I thought that it has covered Asha)

Of the three arguments presented by Keach, two support an agreement analysis (wh-questions and idioms) and one supports a pronominal analysis (habitual clauses). Thus Keach concludes that SM in Swahili is ambiguous between a pronoun and agreement. However, the evidence for a pronominal analysis is ambiguous at best, while there is further evidence supporting an agreement analysis. I present this evidence in the next section.

3. Arguments against a pronominal analysis

In addition to the arguments presented by Keach, there are three further pieces of evidence that I will present that also suggest an agreement analysis. The first two arguments make use of the fact that in a pronominal analysis, SM itself is the true subject and what has traditionally been referred to as subject is actually a topic. I now show that the overt subject in a declarative Swahili sentence does not exhibit topic properties, casting doubt on the pronominal analysis of SM.

3.1 Topics as Quantifiers

One property of topics is that they cannot be quantifiers (Lasnik & Stowell, 1991; Rizzi, 1992):

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

(11) a. Nothing is impossible
b. *Nothing, (it) is impossible
In Nairobi Swahili, this restriction also holds. In (12a), the object (\textit{kila kitabu}) is in object position, and is ungrammatical when topicalized, as in (12b).

\begin{align*}
(12) & \text{a. } & a & \mathbin{-\text{nunu}} a & \text{ kila kitabu} & \text{SA}_{3s} & \text{past} & \text{buy} & \text{IND} & \text{every book} \\
& & & & \text{‘She bought every book’} \\
& \text{b. } & * & \text{ kila kitabu, } a & \mathbin{-\text{(ki)}-\text{nunu}} a & [t] & \text{every thing } & \text{SA}_{3s} & \text{past} & \text{(OA}_{7}) & \text{buy} & \text{IND} [t] \\
& & & & & & \text{‘Every book, she bought’}
\end{align*}

Thus the restriction on quantified topics holds in Nairobi Swahili. Under a pronominal analysis of SM, the preverbal DP is in topic position, and so a quantifier should be ungrammatical. However, quantifiers are possible in preverbal position, suggesting that the preverbal DP is in subject position:

\begin{align*}
(13) & \text{a. } & \text{kila} & \text{ mtoto} & a & \mathbin{-\text{nunu}} a & \text{ ki} \mathbin{-\text{tabu}} & \text{every child } & \text{SA}_{3s} & \text{past} & \text{buy} & \text{IND} & \text{7–book} \\
& & & & & & \text{‘Every child bought a book.’} \\
& \text{b. } & \text{kila} & \text{ ki–tabu} & \text{ki} & \mathbin{-\text{nunuli}} & \text{w} & a & \text{ na} & \text{ mtoto} & \text{every 7–book } & \text{SA}_{7} & \text{past} & \text{buy} & \text{passive} & \text{IND} & \text{by child} \\
& & & & & & \text{‘Every book was bought by a child.’}
\end{align*}

3.2 Topics and Questions

It is generally known that the answer to a question cannot be a topic, as shown in (14), but in Nairobi Swahili, the preverbal DP can be the answer to a question, as shown in (15).

\begin{align*}
(14) & \text{a. } & \text{Who arrived early?} \\
& \text{b. } & \text{?? As for John, he arrived early} \\
& \text{c. } & \text{John arrived early} \\
(15) & \text{a. } & \text{nani} & a & \mathbin{-\text{fik}} & a & \text{ mapema} & \text{who } & \text{SA}_{3s} & \text{past} & \text{arrive} & \text{IND} & \text{early} \\
& & & & & & \text{‘Who arrived early?’} \\
& \text{b. } & \text{?? Juma, } a & \mathbin{-\text{fik}} & a & \text{ mapema} & \text{Juma, } & \text{SA}_{3s} & \text{past} & \text{arrive} & \text{IND} & \text{early} \\
& & & & & & \text{‘Juma, he arrived early.’} \\
& \text{c. } & \text{Juma} & a & \mathbin{-\text{fik}} & a & \text{ mapema} & \text{Juma } & \text{SA}_{3s} & \text{past} & \text{arrive} & \text{IND} & \text{early} \\
& & & & & & \text{‘Juma arrived early.’}
\end{align*}

When the preverbal DP is topicalized (indicated by ‘comma’ intonation) in example (15b), it is awkward as an answer to the question in (15a). Thus topics cannot be the answer to questions in Swahili. In example (15c), the non-topicalized preverbal DP is grammatical as the answer to the wh-question in (15a). This supports the view that the preverbal DP (without ‘comma’ intonation) is not in topic position, but rather in subject position.

3.3 Typological Evidence

A pronominal analysis of SM appears more and more implausible. 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 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, SM has been particularly well-studied in four languages: Luiseño, Cupéño, Serrano, and Cahuilla. In Luiseño, the unmarked word order is shown in (16a) (examples are from Steele, 1995), where the clitic (up) is in second position following the subject (hengeemal):

(16) a. hengeemal up heyiq Subject-clitic-verb
    boy 3sg is:digging
    ‘The boy is digging’

    b. heyiq up hengeemal Verb-clitic-subject
    is:digging 3sg boy
    ‘The boy is digging’

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

In (16a), the unmarked order is subject-clitic-verb. According to Steele (1995, p.1227), (16b) with the verb preceding the clitic is semantically non-distinct from (16a). (16c) – where the clitic sequence is not second – is ungrammatical. This is also true of two of the other three most well-studied languages: Cupéño and Serrano. Thus the order of the clitic and verb is free, provided the clitic is in second position. However, the fourth 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 (prouns)” (Steele, 1995, p.1227).5 In making this distinction, Steele (along with Steele, 1979; 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.

We can now apply this test to the Swahili SM marker to determine whether it is a prefix or a pronominal clitic. The verbal complex in Swahili acts as a unit: when the subject moves, the subject marker remains in its original position, never moving with the subject (examples 17b). Similarly, when the object moves, the verbal complex remains unaffected, with the order of affixes remaining intact (17c).

(17) a. Juma a - na - m - pend - a Mariam
    Juma SA3s-Pres- OA3s-like - IND Mariam
    ‘Juma likes Mariam’

5 The only examples she gives are 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.
b.  
Mariam, Juma  
\begin{align*}
\text{Mariam} & \text{ Juma} \\
\text{SA} & \text{3s} \text{\textsuperscript{-Pres}} \text{ OA} & \text{3s} \text{\textsuperscript{-like}} \text{ IND} \\
\end{align*}

'Juma likes Mariam'

c.  
Mariam, Juma  
\begin{align*}
\text{Mariam} & \text{ Juma} \\
\text{Mariam} & \text{ Juma} \\
\text{SA} & \text{3s} \text{\textsuperscript{-Pres}} \text{ OA} & \text{3s} \text{\textsuperscript{-like}} \text{ IND} \\
\end{align*}

'(as for) Mariam, Juma likes (her)'

In this regard Swahili SM behaves like Cahuilla SM: Word order is fixed with respect to the verb, suggesting that it is an agreement marker rather than a pronominal clitic. I do not contend that this typological fact is convincing by itself. However, in conjunction with the numerous and compelling arguments provided above, it is clear that SM is unambiguously an agreement marker in Nairobi Swahili.

4. Discussion and Conclusion

Summarizing, while a pronominal analysis has been proposed by various authors (e.g., Zwart, 1997), the evidence that SM is a pronoun is weak and unclear. The evidence that SM is agreement, on the other hand, is considerably stronger. The arguments for this latter position include the fact that idiom subjects behave as subjects and not topics, quantifier DPs may occur in subject position (showing this position to be subject position and not a topic position), and that the SM prefix behaves like agreement in other languages, as opposed to a pronominal clitic. Thus I conclude that SM is agreement between the subject and the verb.

The question arises, however, as to why such ambiguity occurs in language at all. If an element is a pronoun, surely it should behave as a pronoun. And if an element is agreement, surely it should behave as agreement. However, the very fact that there is ambiguity in the function and the syntax of these elements suggests a close synchronic connection. Additionally, as has been proposed by several authors, such a close connection may be indicative of a close diachronic connection. For example, Givón (1976) argues that over the course of time pronouns in Bantu languages such as Swahili cliticized to the verb and developed into agreement markers. Thus in a language such as Chichewa, where the function of the SM marker is ambiguous, it would seem that this element is in an intermediate stage between the pronominal stage and the agreement stage.

According to Keach (1995), SM in Swahili is also in an intermediate stage between pronoun and agreement. However, there are several ways to interpret her data. First, it is unclear that all the data that Keach used actually shows ambiguity (as described above in the text). Secondly, the data she uses comes from Zanzibari Swahili – a significantly older, more traditional variety of Swahili. The evidence presented in this paper comes from a Kenyan dialect of Swahili spoken in Nairobi. This variety of Swahili is significantly different from Kiswahili Sanifu (see Deen 2002 for one such difference), and is arguably a creole of Kiswahili KiMvita (spoken in Mombasa), Kikuyu, English, Luo and perhaps other influences. It has developed primarily in the shanty towns of Nairobi, although many speak this variety of language outside of the Shanty towns. Thus a fluid, changing language like Nairobi Swahili is more likely to exhibit innovated features. One way to interpret the evidence presented in this paper is that in newer varieties of Swahili such as Nairobi Swahili, SM has moved through the pronominal stage entirely and is now purely agreement.

One may be skeptical of such rapid change (in the span of one or two generations) given the fact that Bantu languages have remained relatively stable over the course of thousands of years. However, it should be noted that over the history of the Bantu languages, they have arguably never been exposed to the social forces that exist today in East Africa. The linguistic and social diversity that exists today in Kenya and Tanzania in general, but in Nairobi in particular, is unprecedented in the history of Bantu languages. Thus such forces may indeed be responsible for the rapid grammatical change that we now have evidence for.
References


