Hieber, Daniel W. 2008. On the status of the subject agreement marker in Swahili. Term paper, 'Advanced Syntax', Prof. Ann M. Reed, Spring 2008, The College of William & Mary.

ON THE STATUS OF THE SUBJECT AGREEMENT MARKER IN SWAHILI

BY

DANIEL W. HIEBER

ADVANCED SYNTAX

PROFESSOR REED

MAY 7, 2008

Abstract

This paper examines the status of the subject agreement morpheme in modern Swahili. There is a running debate in the literature as to whether this morpheme is a pure agreement marker, an incorporated resumptive pronoun, or something else entirely. A common approach is to suggest that its status lies at some intermediary stage between pronoun and agreement affix. This paper suggests that the relevant halfway point is to analyze the morpheme as a suppletive copula. After citing evidence against both agreement and pronominal analyses, the paper outlines the process by which the subject marker develops from pronoun to copula to agreement marker. This copular approach requires reanalyzing the Swahili verbal paradigm as 'be + participial'. The suggested paradigm is consistent with evidence suggesting that Swahili verbs are in fact highly nominal, so that they function in a similar way to adjectival agreement. Ultimately the paper concludes that the verbal prefixes serve the same function as their morphologically similar cousins, the nominal prefixes.

Contents

§1. Swahili Class Prefixes

Swahili, like most Bantu languages, features a complex system of noun classes, each with its own set of agreement morphemes. Additionally, nouns are marked with prefixes indicating to which class it belongs. While linguists differ in how they organize the various noun classes, we can distinguish sixteen different uses of the class prefixes. Examples of each noun class, with the subject prefixes shown in bold, are given below. Note that the plural of class 11 is *also* class 10 – I note this merely to avoid confusion.

(1) *Swahili Noun Classes* (adapted from Ashton 1959, 10)

1.	**m**-tu	[1-man]	‘person’	2.	**wa**-tu	[2-man]	‘persons’
3.	**m**-ti	[3-tree]	‘tree’	4.	**mi**-ti	[4-tree]	‘trees’
5.	**ji**-cho	[5-eye]	‘eye’	6.	**ma**-cho	[6-eye]	‘eyes’
7.	**ki**-su	[7-knife]	‘knife’	8.	**vi**-su	[8-eye]	‘knives’
9.	**n**-jia	[9-path]	‘path’	10.	**n**-jia	[10-path]	‘paths’
11.	**u**-limi	[11-tongue]	‘tongue’	10.	**n**-dimi	[10-tongue]	‘tongues’
12.	**u**-tu	[12-man]	‘humanity’				
13.	**ku**-cheza	[13-play]	‘playing’				
14.	pahali	[ø(14)-place]	‘place’				
15.	pahali	[ø(15)-place]	‘place’				
16.	pahali	[ø(16)-place]	‘place’				

These noun classes come in pairs of singular and plural, effectively creating a system of eleven ‘genders’, each with its own number morphology. Traditionally each gender is associated with certain semantic classes; for example, classes 3 and 4 are associated with plants, classes 5 and 6 with things which come in pairs or body parts, etc. Classes 14-16 are locative classes (definite, indefinite, and internal locations). Additionally, classes 1 and 2 are exclusively animate, while class 13 is used for infinitivals, and class 12 for some abstract nouns. Note that these class prefixes, like the example for class 12, can be derivational in changing the semantic category of the word. Thus we have the derivations in (2).

(2) *Nominal Derivation*

m-tu	‘person’	m-ti	‘tree’
u-tu	‘humanity’	ki-ti	‘chair’
ki-tu	‘thing’	ji-ti	‘large tree’
ji-tu	‘giant’		

§2. Agreement Markers

Nouns in Swahili has two sets of agreement markers – one for use with adjectives, another for verbal agreement and a variety of other cases. On adjectives, the agreement markers are identical to class prefixes used on the noun:

(3) a. **m**-tu **m**-vivu
1-person 1-lazy

'a lazy person'

b. **wa**-tu **wa**-vivu
2-person 2-lazy
'lazy people'

c. **m**-tu **m**-moja
1-person 1-one
'one person'

d. **wa**-tu **wa**-tatu
2-person 2-three
'three people'

(4) wa-vivu
2-lazy
'lazy (people)'

As Ashton notes (1959, p. 46), this adjectival agreement can be interpreted as a noun in apposition. This is evidenced by the substantive use of the adjective in (3), and indeed in this case the substantive adjective is entirely indistinguishable from other nouns. This is also true for certain quantifiers (cf. (11) below). My point here is that the adjective agreement markers and noun class prefixes are highly related, and so do not require a special analysis to explain their function. I point this out merely to separate the adjective agreement system from the verbal agreement system; my hope is that the former does not fall prey to the same sorts of issues – the issues raised in this paper – which surround the latter.

The other set of agreement markers are used to show both subject and object agreement on the verb, on relative markers, and in a variety of other special cases like demonstratives. Our primary concern here will be verbal agreement, specifically agreement with the subject NP. These verbal agreement morphemes are similar in many respects to the adjectival agreement markers (i.e. the class prefixes), but differ in a substantial number of cases, as is seen in (5) below. The verbal agreement morphemes themselves contain some variation, perhaps justifying their separation into two distinct categories—object and subject agreement. But as the instances of difference are few, and fuzzy besides, I do not find this necessary, especially since we will be dealing primarily with subject marking here.[1]

[1] There are three instances where the object marker differs from the subject marker:

1. The second person plural; S(ubject) M(arker) = m, O(bject) M(arker) = wa. Like English, there are a number of strategies for formulating the second person plural, and they vary by dialect.
2. The third person singular; SM = a, OM = m. In this case, the OM –m– is simply the class prefix.
3. The second person singular; SM = u, OM = ku. This is the one especially contrasting case, to which I can offer no explanation.

It is worth pointing out that all of these cases involve animate personal pronouns, suggesting perhaps a different historical origin than that of the other subject/object markers.

(5) *Swahili Agreement Morphemes*

Class	*Person*	*Agreement Morpheme*	*Class Prefix*
1	1	ni	m
	2	u	
	3	a	
2	1	tu	wa
	2	m	
	3	wa	
3		u	m
4		i	mi
5		li	ji
6		ya	ma
7		ki	ki
8		vi	vi
9		i	n
10		zi	n
11		u	u
12		u	u
13		ku	ku
14		pa	ø
15		ku	ø
16		mu	ø

§3. Uses of the Agreement Morpheme

The most standard use of the agreement morpheme is to mark subject and object agreement on the verb. Swahili verbs are formed using the following (highly simplified) paradigm:

(6) Subject Marker + Tense/Aspect + Relative Marker + Object Marker + Stem + Mood

For example:

(7) M-toto a-me-ki-let-a ki-tabu

1-child Agr1-Perf-Agr7-bring-Ind 7-book[2]
'the child has brought the book'

Here both the subject *mtoto* and the object *kitabu* are marked on the verb. The subject and object NPs are both optional, depending on the focus. One could say simply *amekileta* 'he has brought it', with both subject and object markers referring anaphorically to prior discourse topics. In Standard Swahili (SS), however, the SM is never omissible.

Outside verbal morphology, the agreement morpheme appears in a variety of environments, such as:

a) on demonstrative pronouns

(8) **m**-ti **u**-le
3-tree Agr3-Dist
'that tree'

b) with the relative marker

(9) *miti ambayo ilianguka*
mi-ti amba-**i**-o **i**-li-anguk-a
3-tree Comp-Agr3-Rel Agr3-Pst-fall-Ind
'the tree which fell'

c) on 'possessive' particles

(10) **ma**-ji **y**-a ø-chungwa
6-water Agr6-Poss 5-orange

[2] Abbreviations used in the glosses:
Agr1 (Agr2, etc.) - agreement morpheme, where the number indicates the noun class
1 (2, etc.) - noun class prefix
Perf – perfect tense
Ind – indicative mood
Dist – distal demonstrative
Comp – complementizer
Rel1 (Rel2, etc.) – relative marker
Pst – past tense
Poss – the possessive particle
Pres – present tense
Loc – locative particle
Cop – copula
App – applicative
Dem – demonstrative
Pass – passive

'orange juice' (lit. 'water of orange')

d) with quantifiers

(11) *nyumba zote*
n-umba **zi**-ote
10-house Agr10-all
'all the houses'

e) with possessive pronouns

(12) *mayai yangu*
ma-yai **ya**-angu
6-egg Agr6-my
'my eggs'

f) and finally with the particle *na* 'and/with' (used in conjunction with the agreement morpheme, however, it becomes the verb 'have')

(13) **m**-toto **a**-na ø-baridi
1-child Agr1-with 9-cold
'the child is cold' (lit. 'the child is with cold' or 'the child has cold')

§4. Zwart's Objection

In an article titled 'Rethinking Subject Agreement in Swahili', Jan Wouter Zwart (1997) argues that the agreement morpheme cannot be the subject agreement marker as it is traditionally analyzed. The distribution of the agreement morpheme, Zwart argues, is too broad. It is used in a variety of contexts outside the verbal morphology, as shown above in (8)-(12) (and possibly (13)). On Zwart's definition of a subject agreement marker, i.e. "an affix with no other function than to mark the congruence of the subject and the verb," we are forced to accept the possibility that this agreement morpheme may have some broader function than simple subject agreement (Zwart 1997, 4).

Zwart, however, takes the argument further, and says that this morpheme shows no subject agreement whatsoever. Any concord that exists between this morpheme and a subject NP is not a matter of agreement between a verb and its subject NP, but between the subject NP and something else entirely. That something else, according to Zwart, is a type of resumptive pronoun, i.e. an element or structure which repeats or in some way recapitulates the meaning of a prior element. The resumptive pronoun is not a personal pronoun according to Zwart, but rather something not unlike clitics or discourse-bound demonstratives in other languages.

We will turn to the details of Zwart's argument in a moment, but first it is perhaps elucidating to understand why one might be tempted into such an analysis. Consider the example from Spanish in (14):

(14) yo com-o
I eat-Agr1s
'I am eating'

When one says *yo hablo* as opposed to simply *hablo* in Spanish, it is generally correct to say that the *yo* has an emphatic use, such that the meaning is not simply 'I am eating', as above, but rather something more akin to '*I'm* speaking' or 'as for me, I'm speaking'.[3] In other words, the use of the explicit subject pronoun in conjunction with the subject agreement morpheme on the verb provides for a contrastive focus. Obviously, this depends on the status of the discourse topic: if we have already been talking about Carlos, then to say *Carlos come* 'Carlos is eating' emphasizes the fact that *Carlos* is the one eating; if this is the first time that Carlos has been mentioned, then we would not expect this contrastive focus.

The difference is this: it would appear that verbal subject agreement refers anaphorically to the topic of the discourse in all cases, so that inserting an overt subject NP stands in contrast to the preestablished discourse topic. When the discourse topic *is* the overt subject NP, however, the subject agreement marker refers back to *it*, so that no semantic contrast occurs. It follows that the insertion of a first- or second-person pronoun in the subject NP will *always* be contrastive, because the discourse topic ('I' or 'you') is clear on the basis of the agreement morpheme alone. Third-person subjects, however, require a named referent, and so can be either contrastive or not.

Swahili functions similarly. Example (15) has a contrastive meaning, whereas simply saying *ninaenda* does not.

(15) mimi ni-na-end-a
I Agr1-Pres-go-Ind
'*me*, I am going'

The semantic interpretation of both (14) and (15) show a kind of double pronoun where one resumes the other. On this analysis it is easy to understand why we might want to consider the subject agreement marker as some kind of resumptive pronoun.

There is other evidence for this view. First is the simple fact that typologically, verbal agreement does stem historically from the grammaticalization of pronouns onto the verb (Givón 1976). This should immediately make us cautious; the status of the agreement morpheme in Swahili could lie anywhere on a diachronic continuum from a pronominal analysis to one of agreement, so this morpheme may not fall clearly into any one category. Bresnan & Mchombo (1987) in fact claim that the agreement markers in Chichewa (a Bantu language with similar patterns of verbal agreement to Swahili) remain ambiguous between a pronominal analysis and an agreement one, and have properties of both.

Unlike Swahili, however, Spanish subject agreement is true subject agreement, in that the agreement morphology is found only in verbal constructions, and used exclusively to mark agreement. As Zwart has shown us, in Swahili this is not the case.

§5. Agreement versus Pronoun[4]

It is important here to clarify what is meant by 'agreement analysis' (the standard interpretation) and 'pronominal analysis' (Zwart's analysis). An agreement analysis is one which assumes that the subject raises from Spec-VP to Spec-AgrP, and that this spec-head relationship

[3] What we mean by 'emphatic' is admittedly hazy. Here it is perhaps best understood as a Topic with contrastive focus.

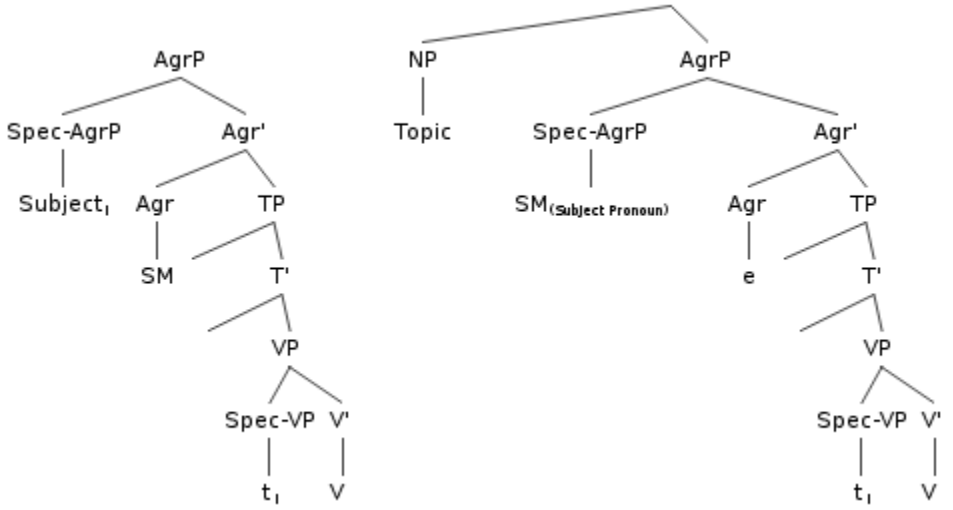
[4] This section is heavily modeled on §1.1 of Deen's (2006) paper 'Subject Agreement in Nairobi Swahili', where his purpose is the same – to better explicate the views of the opposing camps.

triggers agreement. The SM will fill the Agr head, and the subject with which it agrees is in the specifier. A pronominal analysis assumes that the SM is a pronominal DP raised from Spec-VP to Spec-AgrP, leaving the Agr head empty or phonetically null.

The crucial difference between the two is that by a pronominal analysis, in sentences with overt 'subject' NPs, that NP is in fact a Topic that binds the pronominal clitic in subject position. (16) compares these two analyses.

(16) *Agreement v. Pronoun*

<u>Agreement</u> <u>Pronoun</u>



An implication of this is that, under a pronominal analysis, a preverbal subject NP in Swahili will have the properties of a Topic, and not a subject. These facts are important to keep in mind as we tackle Zwart's argument.

§6. The Argument against Agreement Affixes

If Zwart is to prove his arguments, he must do two things: first, he must show that the agreement morpheme cannot possibly be a subject agreement marker. Secondly, he must motivate the analysis of the agreement morpheme as a resumptive pronoun. Both are necessary if Zwart's claim that Swahili verbs take no subject agreement is true. We turn now to the first task.

Zwart has already given one fairly convincing argument against an agreement analysis, namely that the distribution of the agreement morpheme is far too broad. His second argument has to do with agreement *ad sensum*. In Swahili, many nouns for animals belong to one noun class yet take the agreement morpheme of another. For example (taken from Zwart 1997):

(17) **wa**-le **vi**-jana **wa**-na-chez-a m-pira
Agr2-Dist 8-young Agr2-Pres-play-Ind 3-ball
'those youngsters are playing ball'

Here, 'youngsters' is a class 8 noun which takes class 2 agreement morphology. This occurs because *vijana* is an animate noun semantically, and classes 1 and 2 are associated semantically with animates.

The problem this poses for a subject agreement analysis is that the verb agrees *semantically*, not syntactically, with the noun *vijana*. This is not typical of subject-verb agreement. Subject agreement marks agreement between the *noun class* and the verb, *not* the semantics and the verb. At the same time, Zwart uses this as support for his resumptive pronoun analysis: this kind of agreement *ad sensum* isn't unusual in constructions where a pronominal element resumes a previously mentioned entity. He uses an example from Dutch to illustrate this:

(18)

Dat	meisje	**die**	is	gek
Dist-Ntr	girl	Dist-Nntr	is	crazy[5]

'that girl is crazy'

I however question whether this limitation on the function of subject agreement is correct. Even in English subject-verb agreement we sometimes see semantic agreement, or agreement *ad sensum*. Consider the examples in (19):

(19) a. The committee *is* meeting today. (American English)
b. The committee *are* meeting today. (British English)
c. This set of china *is* used only on formal occasions, and *are* all hand-painted.

(19c) is particularly interesting, as it conjoins verbs with different agreement, *is* and *are*. To me this suggests both a grammatical agreement (*is*) and a semantic one (*are*) in simultaneous use. So I am not sure that this second argument is entirely convincing.

Zwart's final argument concerns existential constructions in Swahili. (20a) is an example of this type of construction. *–po* here consists of the agreement morpheme for class 14 (the class used for definite locatives), plus what is traditionally known as the '–o of reference' (Ashton 1944, 18).

(20) a. *wapo*
wa-pa-o
Agr2-Loc
they-here
'here they are'

b. wa-li-po
Agr2-Cop-Loc
they-are-here
'here they are'

Before tackling Zwart's argument, some historical background is necessary. McWhorter (1992) tells us that (20a) originally developed from the construction in (20b). In early modern Swahili, the language had a full verb *–li*, meaning 'be', which took the agreement prefixes in

[5] Neuter (Ntr) and nonneuter (Nntr) are the only two genders in Dutch.

(21), forming constructions like that in (22). Some reflexes of this copula still exist in certain relative constructions today (e.g. *u-li-ye* 'you who are'):[6]

(21) *Agreement Prefixes (FSPs)*

	Singular	Plural
1	ni	tu
2	u	m
3	a/yu[7]	wa

(22) Hamisi yu-li m-pishi
H. Agr1-Cop 1-cook
'Hamisi is a cook'

(23) wewe u m-kurugenzi
you you 1-director
'you are the director'

Note that these agreement prefixes are the same as in SS today (cf. (5) above). At the same time, Swahili had another copula strategy, shown in (23), which was simply the use of what are known as the Free Subject Prefixes (FSPs; the agreement prefixes just listed) in place of the copula. McWhorter (1992) outlines the process by which *–li* disappeared and was replaced with the FSPs.

One theory, put forth by Perrott (1951), is that the FSPs were the leftover result of the erosion of full forms of *–li*. Constructions like (22) in time would have eroded to just (24). According to the theory, this agreement affix would have been reanalyzed as a copula.

(24) Hamisi yu m-pishi
?H. Agr1 1-cook
?'Hamisi is a cook'

McWhorter offers evidence against this, mainly that a) we do not find intermediate forms between *yuli* and *yu*, such as *yul'*; and b) both the full form of *–li* and the construction using FSPs co-occur as copula strategies in some of the earliest documents surveyed. I also point out the potential ungrammaticality of (24).

Instead what happened was that *–li* was reanalyzed over the course of the 18th and 19th centuries as a past tense affix, and the strategy in (23) rose to prominence. As the resumptive pronoun in (23) occupies the copular slot in the grammar, it came to be analyzed as the copula. At first this would seem favorable to Zwart's analysis – if the SM was originally a resumptive pronoun, then it is not much of a leap to call the SM a cliticized pronoun. Later, however, I will show how the resumptive pronoun developed into a copula, and was then incorporated into the verbal paradigm. Zwart's argument, I believe, is based on historical traces of an outdated form. Under this interpretation, the historical origin of the SM is in fact most favorable to a copular analysis, rather than to a pronominal one.

[6] The following examples are taken from McWhorter (1992).
[7] *Yu-* alternates with *a-* both dialectally and historically. Today, *a-* is close to taking over the last uses of *yu-*.

But to return to the argument: as a result of this reanalysis from pronoun to copula, the *–li* morpheme disappeared from locative constructions, so that in SS today, locatives are formed as in (25), necessitating an empty copula.

(25) wa-ø-po
Agr2-Cop-Loc
they-are-here
'here they are'

Zwart's argument is this: if (25) is in fact the correct analysis of Swahili locatives (later I will question this), then *wa-* cannot be a subject agreement affix, because typologically such affixes do not attach to a zero verb. His solution to this is of course to suggest that *wa-* is a pronoun of some kind, meaning that the structure is no longer anomalous. In contrast to this, most linguists today will analyze *wapo* as a locative copula, where *–po* is the copular element, and *wa-* is the SM.

My own take on the matter is somewhat different. We know that the SM developed from a copula, and that the locative *–po* is a combination of the locative nominal prefix *pu-* and the 'o-of reference'. Therefore it makes more sense to suggest that *wa-* still serves some verbal role, resolving the anomaly pointed out by Zwart, while retaining a copular element in accordance with the analysis put forth by most linguists.

So Zwart has three main arguments against the analysis of the agreement morpheme as a type of subject-verb agreement: a) the distribution doesn't match that expected for a subject agreement marker, b) Swahili agreement *ad sensum* does not fit normal patterns of subject-verb agreement, and c) the agreement morpheme in locatives cannot be an SM because it would have to attach to a zero verb. The first and third of these I find compelling, while the second seems to me still open to debate. But they are only compelling insofar as they disprove a 'pure' agreement analysis; they do not, however, motivate a pronominal analysis in and of themselves.

In light of this evidence, I am inclined to agree with Zwart in saying that the agreement morpheme cannot be just a marker of subject-verb agreement. Now we turn to Zwart's stronger claim, that this agreement morpheme is in fact a resumptive pronoun. If correct, this implies that the agreement morpheme cannot be a marker of subject-verb agreement at all (as opposed to being a pronoun which *also* marks for subject-verb agreement, which would be unusual).

§7. The Argument for a Resumptive Pronoun

Here I will examine each of Zwart's arguments for the resumptive pronoun analysis, raising objections to each in turn. In §8 and §9 I will raise some independent objections and attempt to show that this analysis is untenable, and in §10 I proceed to offer my alternative – a copular analysis.

§7.1 Pronominal Distribution

First, Zwart correctly states that the distribution of this morpheme is close to that of a pronoun. As Crystal tells us, "The grammatical statement of pronominal distribution in a language is usually quite complex" (Crystal 2003). Likewise, the distribution of our yet-to-be-determined agreement morpheme is varied and complex. Yet one could say the same of many linguistic entities (consider the use of the possessive *–a* particle in Swahili, which functions as far more than simply a marker of possession, and has a broad distribution). So while this matter

of distribution might possibly contribute to a corpus of converging evidence for a resumptive pronoun analysis, it is not conclusive in itself; it is useful only as supporting data.

§7.2 Symmetry with Object Marking

Second, it is generally accepted (following Bresnan & Mchombo) that the object marker (OM) in Bantu languages is not a grammatical agreement marker, but rather an incorporated object pronoun. In his analysis of the SM, Zwart hopes to make an analogous claim. Meanwhile, Brsenan & Mchombo (1987) and Keach (1995) claim instead that the SM is ambiguous between a grammatical agreement marker and an incorporated pronoun.

There is an immediate appeal to theoretical symmetry in attributing analogous functions to both the SM and the OM, but we should not let that persuade us. Languages are complex creatures, after all. We might think that asymmetries between the SM and OM point to different functional roles. For example, object marking is optional depending on whether or not the object is the Topic; the OM is obligatory when this is the case, and optional otherwise. Opposite to this, the SM is always obligatory, regardless of the discourse function of the subject. Note also that it is typologically unusual for subject agreement to be optional, while with pronouns this is usual, providing more evidence for a pronominal analysis.

Zwart would resolve this subject-object asymmetry, I think, by saying that the SM appears in all contexts because the subject NP is always the Topic (this follows from his pronominal analysis), whereas the object NP is not. Thus the asymmetry is resolved. If there are no asymmetries, then, we are still left with the question of what motivates treating the subject NPs and object NPs differently. It would seem odd for a verbal system to use a single set of morphemes to mark grammatical agreement with one NP (the subject) while simultaneously marking anaphoric agreement to another (the object). If we are to disagree with Zwart, we should be able to explain why 'agreement' (in the broad sense of either grammatical or anaphoric) is not uniform even within the structure of the clause. Deen (2006) tries to address this problem simply by noting that Swahili – like Chichewa – may still be in a state of transition between a pronominal analysis and an agreement one (a diachronic ambiguity mirrored by a functional one in the analysis of the morpheme), but is primarily a case of grammatical agreement.

There is evidence that, at least in Swahili, the ambiguity between a pronominal analysis and a grammatical agreement analysis is disappearing. Deen suggests that the newest forms of Swahili, such as those found around Nairobi, may have already moved through the pronominal stage and now show full grammatical agreement with both the subject and object, even though other Bantu languages have not yet seen this change. This rapid change is likely due to Swahili's status as a lingua franca in the region, as well as a variety of new social pressures unprecedented in the history of East Africa (Deen 2006, 232).

Even as they establish the pronominal status of the OM in Chichewa, Bresnan & Mchombo also note that Swahili might serve as an exception to the pronominal analysis. One of the predicted results of their pronominal analysis is that it should not be possible to question the overt NP and still retain the marker on the verb. But obviously the overt subject NP cooccurs with the SM, and – unlike Chichewa – in Swahili it is also possible to question the overt object NP while retaining the OM. The two results are contrasted in (26) (Bresnan & Mchombo's number (41)) and (27) below, with the potentially offensive OM highlighted in bold.[8]

[8] Examples (27) and (28) are taken from Bokamba (1981) as cited in Bresnan & Mchombo (1987, pp. 760, 777).

(26) *Chichewâ*

??(kodí)	mu-ku-***chí**-fún-á	chí-yâni?
Q	you-Pres-Agr7-want-Ind	7-what

'what do you want (*it)?'

(27) *Swahili*

Bakari	a-na-**wa**-som-e-a	nani	hadithi	maktaba-ni?
B.	Agr1-Pres-Agr2-read-App-Ind	who	stories	library-Loc

'to/for whom is Bakari reading stories at the library?'

As Swahili allows for the presence of the OM while the object NP is questioned in place, this suggests that the OM is in fact *not* an incorporated pronoun in Swahili, but has achieved full status as a marker of grammatical agreement in Swahili.

If we accept this as evidence for an agreement analysis, then obviously Zwart's plan to draw the analogy between subject and object marking – thus proving his pronominal analysis – fails in its objective. Of course this data gets hazy, and different dialects of Swahili likely weigh in differently on the issue. Whatever the case, though, Bresnan & Mchombo's analysis with its implications regarding Focus, Topic, and anaphoric agreement between incorporated pronouns may not be as relevant to Swahili as first thought.

However I do not believe this settles the matter. Zwart still presents good evidence to the view that the SM cannot be simply an agreement marker. (Likewise, the optionality of the OM still suggests that it cannot be a pure agreement marker either.) It seems there must be a third option, and one which (I think) avoids the need to retain a symmetry between object marking and subject marking. Historical evidence suggests that before the SM became an agreement morpheme in the dialects where it has, it was first reanalyzed as a copula. This would appear to be the elusive 'middle ground' between a pronominal analysis and an agreement one. Furthermore, analyzing the SM as a suppletive copula would change the structural paradigm of the verb, so that we would no longer expect a symmetry between the SM and the OM (more details on this later).

§7.3 Locative Inversion

Zwart's third argument for the resumptive pronoun is that in certain locative constructions, the behavior of the SM can be explained by using a resumptive pronoun. The construction at issue is what Zwart calls locative inversion. It is unique because the verb seems to take agreement with a fronted locative, rather than with what would normally be the subject NP (the logical or thematic subject). Example (28) (Ashton 1944, 127-128) gives examples of this construction in use.

(28) a.

mw-itu-ni	m-me-lal-a	wa-nyama
3-forest-Loc(16)	Agr16-Perf-sleep-Ind	2-animals

'in the forest there have slept animals'

b.

ha-pa	pa-me-kuf-a	simba
Dem-14	Agr14-Perf-die-Ind	lion

'here there has died a lion'

c. *kule mjini kumekufa watu wengi*

ku-le	mji-ni	ku-me-kuf-a	wa-tu	wa-engi
15-Dem	town-Loc(15)	Agr15-Perf-die-Ind	2-person	2-many

'there in the town there have died many people'

Zwart (p. 11) notes the strange manner in which these verbs take agreement, and contrasts it with Dutch (while I offer the example from English):

(29) in de tuin zitt-en mensen
in the garden sit-Pl people
'in the garden there sit people'

(30) Here lies Lulu.

In both cases, the verb shows agreement with the postverbal subject NP, and not the preverbal locative NP. Assuming that the analysis of Dutch and English should serve as our model (i.e. that the verb should show agreement with the logical subject NP and not the locative NP), Zwart puzzles at the examples in (28). Why are they different from Dutch and English in their agreement? According to Zwart, the Swahili examples are instances of Topicalization, and so the SM must be analyzed as a subject pronoun, with the locative NP serving as the Topic.

A note on terminology before continuing: as many different linguists have contributed to this discussion in the literature, there is a great deal of equivocation going on with the terms being used. Though we might wish otherwise, 'subject' is a very loose term, and can refer to any of the following: the logical subject, the grammatical subject, the case-assigned subject, the preverbal subject NP, or the subject as defined by its position in the spec-head relationship with the verb. To the best of my ability, then, I will try to keep the following definitions for 'subject' and 'logical subject' for Swahili:

(31) Subject = Df. in a spec-head relationship between Spec-Agr and Agr°, the NP which fills Spec-Agr, thus serving as the specifier to the verb.

(32) Logical subject = Df. in a sentence, the NP which has the 'underlying' thematic role of Agent (i.e. the NP which at the level of the D-structure fills Spec-VP).

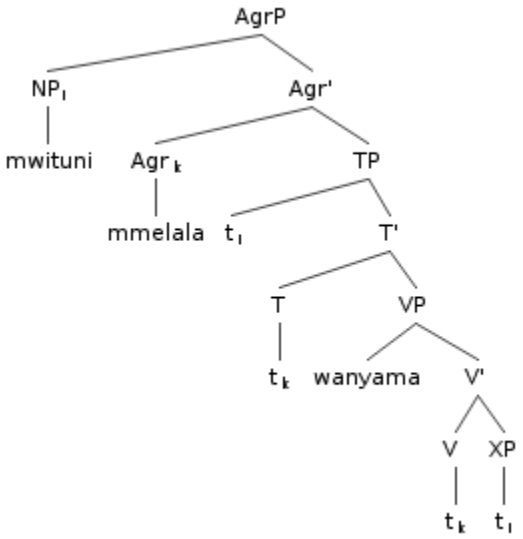
I believe these two definitions will adequately encompass the range of concepts necessary for this discussion. Notice that (31) includes the terms 'grammatical subject' (because grammatical agreement occurs with whatever lies in Spec-Agr), and possibly 'case-assigned subject', if we assume that nominative case is also assigned in Spec-Agr. I will attempt to avoid terms such as 'overt (subject) NP' and 'preverbal subject', unless the ambiguity is intended.

An initial problem for Zwart's analysis is this: if these examples are in fact cases of topicalization, and Zwart's resumptive pronoun analysis is right, then we expect the locative NP to have the Topic role, and the SM/resumptive pronoun to have the subject role. But from a syntactic analysis of this type of construction, we know this is not the case.

Locative inversion is an example of a more general phenomenon in Swahili (and other Bantu languages) known as subject inversion (subject here referring to the logical subject of the sentence). As Demuth & Harford tell us (1999, p. 34), "Bantu languages seem to differ from languages like English in requiring the Topic to be the grammatical subject of the sentence rather than preferring the Agent in that position." So the Swahili SM does not mark agreement with the logical subject, but rather with the Topic, even though they can coincide. At first this seems to confirm Zwart's view – if the fronted NP is always the Topic, and the Topic is sometimes distinct from the subject in Swahili, then it is at least consistent to suggest that the SM might be a pronominal subject. But if we can show that the subject role is not assigned to the SM, then we will have disproved Zwart's hypothesis.

The syntactic tree for cases of locative inversion like (28) is shown in (33) below. Locative/subject inversion occurs whenever an argument other than the logical subject (e.g. an object, a locative) is raised to Spec-AgrP.[9] This blocks the movement of the VP-internal subject into Spec-AgrP, as the node cannot be doubly filled. The matrix verb then raises into Agr to pick up agreement, stranding the logical subject in Spec-VP.[10] That stranding is what gives rise to the apparent subject inversion. We also know that this logical subject must reside in Spec-VP because of the ungrammaticality of sentences like (34).

(33) *mwituni mmelala wanyama*



(34) *Maseru ho-ile kajeno ba-sadi (Demuth and Harford 1999, 7)
Maseru Agr17-go/Perf today 2-woman
'to Maseru they went today, the women'

Example (34) is taken from Sesotho, another Bantu language with subject inversion. Cases like this do not permit any intervening material (here the adverb *kajeno* 'today') between the verb

[9] I am working under the following assumptions: a) the subject is generated in Spec-VP, b) the verb raises to Agr to obtain agreement, c) the subject receives nominative case by raising to Spec-AgrP, and d) Spec-Head agreement occurs between the verb in Agr and the NP in Spec-AgrP.
[10] Verb-raising is a common feature of Bantu languages.

and the logical subject, providing evidence that the logical subject is in fact in the Spec-VP. This restriction holds for Swahili as well, but the point here in using Sesotho is to illustrate that these facts hold cross-linguistically for Bantu languages as well.

Grammatical agreement between the Topic NP and the verb is triggered by the Spec-Head relationship in the AgrP. Logical subjects that remain in the Spec-VP obviously do not trigger such agreement, and in fact such agreement is ungrammatical, as shown by the Sesotho example in (35) and the Swahili in (36):

(35) *Sesotho*

*Maseru	ba-ile	ba-sadi	kajeno	(Demuth and Harford 1999, 7)
Maseru	Agr2-go/Perf	2-woman	today	

'to Maseru women went today'

(36) *Swahili*

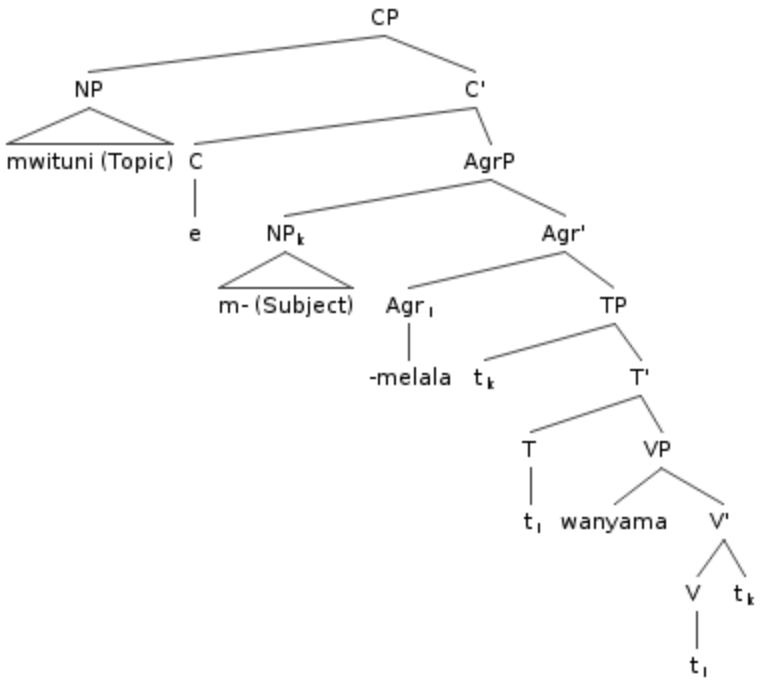
*Nairobi	wa-li-enda	wa-toto	leo
Nairobi	Agr2-Pst-go	2-child	today

'to Nairobi children went today'

Under this analysis of subject inversion, if we assume that the subject receives nominative case in Spec-AgrP, then Zwart's analysis does not hold. If we know that a) the fronted Topic lands in the Spec-AgrP (implied by Demuth, quoted above), and that b) Spec-AgrP assigns nominative case (i.e. the subject role), then the SM cannot also be the subject, as Zwart would like. This seems to refute his analysis of the SM as a resumptive pronoun.

Zwart would respond to this by denying (a) above. What lands in the Spec-AgrP and receives the role of subject is the SM, not the Topic NP. The fronted Topic NP thus lands above AgrP, presumably in CP. But it is not clear how such a tree would work. Where does the resumptive pronoun originate? Does it generate in the complement of the verb? But perhaps the beginnings of Zwart's analysis can be shown in (37) below:

(37) *?mwituni mmelala wanyama*



Alternatively, we might reply to Zwart by claiming – as Machobane (1987) does – that the logical subject, and not the fronted NP, receives nominative case from the Spec-VP. As we cannot be certain, this does not seem sufficient enough to refute Zwart's hypothesis.

We might then try a different approach. One result of Zwart's analysis is that the resumptive pronoun in locative inversion constructions like those we've been seeing is functionally equivalent to the demonstrative in (38) (an example he gives from Dutch) below:

(38)

in de tuin	daar	zitt-en	mensen
in the garden	Dist-Loc	sit-Pl	people

?'in the garden there, there sit people'

Yet Swahili also allows for demonstrative pronouns in such constructions, as in (28c), repeated here:

(28)

ku-le	mji-ni	ku-me-kuf-a	wa-tu	wa-engi
15-Dem	town-Loc(15)	Agr15-Perf-die-Ind	2-person	2-many

'there in the town there have died many people'

Clearly here the demonstrative pronouns of each language serve the same functional role in the sentence. Since the Dutch pronoun is already equivalent to the Swahili pronoun, it cannot possibly be equivalent to the Swahili SM.

§7.4 Quasi-Passives

While the above objections relate specifically to locative inversion, this is really just a part of a broader issue involving the fronting of NPs in Swahili sentences. Postverbal locative NPs in Swahili are interesting because they seem to function much like any other postverbal object. Most notably, they can become passivized in exactly the same manner as the Theme or Patient, as shown in (39a, b). Similarly, either the locative NP or the object NP can raise in subject inversion constructions, and we call this locative inversion or quasi-passive, respectively. An example of the quasi-passive is given in (39c).

(39) a. *Passivized Theme*
Juma a-li-wek-a ki-tabu meza-ni
J. Agr1-Pst-put-Ind 7-book table-Loc
'Juma put the book on the table'

b. *Passivized Locative*
meza-ni pa-li-wek-w-a ki-tabu na Juma
table-Loc Agr14-Pst-put-Pass-Ind 7-book by J.
'on the table was put a book by Juma'

c. *Quasi-passive*
vy-akula vi-li-kul-a wa-toto
8-food Agr8-Pst-eat-Ind 2-child
'the children ate the food'

The active equivalent of (39c) would be *watoto walikula vyakula* 'the children ate the food'. Compare (39c) with (28a) *mwituni mmelala wanyama*, and you can see that the constructions are indeed the same.

The traditional syntactic analysis of the Swahili quasi-passive is to view it as a case of subject inversion, just as outlined for locative inversion above (briefly, an object NP raises to Spec-Agr, blocking NP movement out of Spec-VP and stranding the logical subject). Again, Zwart claims that the fronted NP is a Topic, and does not lie in a spec-head relationship with the verb (meaning that it cannot be the grammatical subject).

First a few terminological items: the traditional syntactic analysis of subject inversion involves A-movement, i.e. movement to an argument position (here, Spec-Agr). As Zwart denies that the overt NP lands in the Spec-Agr, he cannot call this A-movement. Instead, Zwart breaks with the traditional analysis and calls the quasi-passive an instance of fronting, of which there are two types: topicalization and left-dislocation.[11] In left-dislocation, the fronted NP is typically offset by commas, and requires the presence of the OM on the verb in Swahili. It also does not trigger agreement morphology on the verb (indicating that it does not lie in Spec-Agr, and that this is not a case of subject-verb inversion). Thus the fronted NP is likely the Focus of the sentence. (40a) is an example of such left-dislocation. In topicalization, the fronted NP is not offset by commas, does take agreement morphology, and does not allow an OM on the verb.

By analogy to left-dislocation, Zwart claims, it is likely that topicalized NPs too do not lie in Spec-Agr. If quasi-passives are indeed a case of topicalization as Zwart suggests, this implies

[11] Zwart gets this distinction from Bokamba (1987).

that the preverbal NP is a Topic, but not the subject. This is compatible with his resumptive pronoun analysis.

Zwart notes that these two types of fronting exist in Dutch as well, as seen in (40a, b), and that the topicalization construction in (40b) a) mirrors Swahili topicalization in that it also does not allow for the presence of an object pronoun, and b) it includes an (albeit optional) resumptive pronoun. By analogy, he argues that Swahili topicalization involves such a pronoun as well.

(40) a. *Left-dislocation*

Dat boek,	ik	heb	het	nog nooit gezien
that book,	I	have	it	yet never seen

'I have never seen that book before'

b. *Topicalization*

Dat boek	**dat**	heb	ik	nog nooit gezien
Dist-NTR book	Dist-NTR	have	I	yet never seen

'I have never seen that book before'

This argument hinges critically on whether Zwart is right in calling the quasi-passive a case of topicalization, rather than A-movement, and I think he is wrong to do so. His argument goes as follows: the quasi-passive cannot involve raising-to-subject (A-movement) because it is itself derived from another passive form. We know this from the behavior of ditransitive verbs in Swahili. In ditransitive constructions, only the indirect object may be passivized using regular passive morphology, as in (41a). Passivizing the direct object, as in (41b) is ungrammatical. It is however possible to passivize the direct object in quasi-passives like (41c). According to Zwart, (41c) can only be derived from (41b).[12]

(41) a.

wa-toto	wa-li-pik-i-w-a	ch-akula	na	m-vulana
2-child	Agr2-Pst-cook-App-Ind	7-food	with	1-boy

'the children were cooked for by the boy'

b.

*ch-akula	ki-li-wa-pik-i-w-a	wa-toto	na	m-vulana
7-food	Agr7-Pst-Agr2-cook-App-Pass-Ind	2-child	with	1-boy

'the food was cooked for the children by the boy'

c.

ch-akula	ki-li-pik-i-w-a	wa-toto	na	m-vulana
7-food	Agr7-Pst-cook-App-Pass-Ind	2-child	with	1-boy

'the food was cooked for the children by the boy'

The problem is, (41b) is also a passive construction. And if the quasi-passive is derived from another passive, it cannot be A-movement which derives it, namely because the Spec-Agr position is already filled in the underlying form. If the quasi-passive is not a case of A-movement, then it is decidedly a case of topicalization, and thus involves a resumptive pronoun.

But we can easily escape this problem, along with the resumptive pronoun analysis, simply by making a few concessions. If we accept Zwart's point that quasi-passives can have passives as

[12] In contrast, I would claim that, following our model in (39), (41c) is derived from (41a) instead. Unfortunately, the end result in this line of argument will be the same, so it does little good to pursue the point further.

their underlying form, and accept that the filled specifier position in the regular passive blocks A-movement, we might still suggest a *second* specifier position – as Chomsky has done in the past (Categories and Transformations 1995) – into which the fronted NP can raise. The multiple specifier proposal is nice because it allows for a neat parallel in ditransitives, where we might expect an agreement projection for each postverbal object, and expect for each of those agreement projections to have a specifier position that can be filled. Zwart thinks this is problematic because it would wrongly predict Bantu languages to allow optional passivization of either the direct or indirect object. Yet we know from Bresnan & Moshi (1990) that this restriction on the passivization of objects stems from independent grounds; namely, what they call the Asymmetrical Object Parameter (AOP), which states (roughly) that for asymmetrical object languages (of which Swahili is one), only one argument role can be classified as unrestricted, i.e. can function as a 'primary' object, one of whose properties is to allow for passivization. So while the multiple specifier proposal overgenerates, the AOP restricts Swahili's grammatical sentences to those which passive the indirect, and not the direct, object. And if we allow for a tree with multiple specifiers, then it is possible to derive a quasi-passive like (41c) from (41b) and still have *grammatical* agreement (i.e. a spec-head relationship), thus avoiding a resumptive pronoun analysis.

§8. Objections to a Pronominal Analysis

Now that we have raised objections to each of Zwart's positive arguments, I would like to present several independent objections to the resumptive pronoun analysis.[13] This section will consist of a brief overview of objections raised elsewhere in the literature, while the following section discusses one of the major implications of Zwart's analysis and why we should object to it. This will in turn lead us to my own analysis of the SM as neither agreement marker nor pronoun, but rather a suppletive copula. I will also try to draw out support for my analysis here.

§8.1 Swahili Pronouns

One initial and potentially major objection to Zwart's analysis is that, if the SM is a pronoun of some kind, it cannot be a *personal* pronoun. He himself admits to this (p. 5), and the evidence comes simply from the morphology. The personal pronouns, listed in (42) below, are clearly reduplicated forms, as evidenced by their non-reduplicated occurrence in some contexts (specifically, in combination with the preposition *na* 'and/with'). There does not seem to be any relation between the personal pronouns in (49) and the SM in (22).

(42) *Swahili Pronouns*

Number	**Person**	**Free Pronoun**	**Bound Form**	**SM**
Singular	1	mimi	na-mi	ni
	2	wewe	na-we	u
	3	yeye	na-ye	a
Plural	1	sisi	na-si	tu
	2	nyinyi	na-nyi	m
	3	wao	na-o	wa

[13] Parts of this section – including examples – draw heavily from Deen (2006).

This may be problematic for Zwart because it is clear that the SM – whatever its current status – arose historically from a personal pronoun (as outlined in Givón (1976)). Its original use was as a resumptive pronoun as in (24) (repeated here as (43)),

(43) wewe u m-kurugenzi
you you 1-director
'you are the director'

but it was later reanalyzed as a copula, and then became cliticized into the verbal morphology. Assuming that Zwart wishes to draw the analogy to Dutch by calling the modern SM in Swahili some type of demonstrative, we note that nowhere in this process does the pronoun take on a demonstrative-like status. Nor would it make sense to say that the SM obtained some sort of demonstratively pronominal status *after* the process of cliticization had begun.

Also, in calling the SM a resumptive pronoun, are we also to expect pronominal behavior on the part of the SM morpheme in all its various uses (listed in §3 above)? This seems unlikely. The same broad distribution which leads us to think that the SM is more than a simple agreement affix should also make us question its pronominal status. In fact, what we should realize is that, whatever our analysis of the SM when used in the verbal complex, it may simply be that this analysis will not hold for all the varied appearances of this morpheme in the language; in which case we ought not to give as much credence to Zwart's original motivations for doubting the agreement status of the SM. It is worth noting, however, that a copular analysis of the SM *does* seem to hold in a larger variety of the SM's uses.

§8.2 Defining Resumptive Pronouns

Second we note that, per definition, the term 'resumptive pronoun' does not seem fully applicable here. According to Crystal (2003, p. 376), we should expect the following from a resumptive pronoun: a) that there exists a prior lexical element that is the antecedent of the SM, b) that the SM itself has a lexical (not just functional) status, and c) that the SM and its antecedent share the same meaning (or perhaps referent). It is not clear that these all hold, especially (b), since the SM long ago lost any lexical meaning it had upon acquiring copula status. Also, the overt NP cannot always serve as the antecedent to the SM, specifically in constructions where the logical subject occurs postverbally but still triggers grammatical agreement:

(44) a. a-me-enda yeye (Vitale 1981)
Agr1-Perf-go he
'he has gone'

b. vi-tu a-li-vyo-vi-fanya Juma ni vi-baya sana
8-thing Agr1-Pst-Rel8-Agr8-do J. Cop 8-bad very
'the things which Juma did are very bad'

Constructions such as (44) are not typically analyzed as any type of exceptional topicalization or movement, but rather as simplex sentences. This seems to rule out the possibility of anaphoric agreement between the SM and the NP with which it agrees.

§8.3 Topics & Questions

A third strong argument against a pronominal analysis is given by Keach (1995) utilizing principles from Bresnan & Mchombo (1987), relating to Wh-questions in Swahili. According to Bresnan & Mchombo, questioned constituents always bear the Focus function, and an argument cannot bear both the Topic and Focus functions in the same clause. In Swahili, the Wh-word occurs *in situ* as follows:

(45) $nani_i$ a_i-me-enda?
who Agr1-Perf-go
'who has gone?'

Under Zwart's analysis, because *nani* occurs *in situ*, it should bear the Topic function, with the SM serving as the subject. If this were so, *nani* would then bear both Topic and Focus functions, which is explicitly prohibited by the principles posited by Bresnan & Mchombo.

Another property of questions is that they cannot be answered with a Topic. Consider the semantic oddness of (46b).

(46) a. Who arrived early?
b. ?? As for John, he arrived early
c. John arrived early

(47) a. nani a-li-fika mapema
who Agr1-Pst-arrive early
'who arrived early?'

b. ??Juma, a-li-fika mapema
J. Agr1-Pst-arrive early
'Juma, he arrived early'

c. Juma a-li-fika mapema
J. Agr1-Pst-arrive early
'Juma arrived early'

This is true of Swahili as well, as exemplified by the topicalized response in (47b), mirroring (46b). If we analyze (47c) as another case of Topicalization – as Zwart wishes – then this too should be semantically odd, but this is not the case.

§8.4 Topics as Quantifiers

Lasnik & Stowell (1991), as well as Rizzi (1992), point out that Topics cannot be quantifiers. (48) and (49) illustrate this principle in English.

(48) a. I did everything
b. *Everthing, I did (it)

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

This also holds for Swahili, so that (50b) is ungrammatical.

(50) a. a-li-nunua kila ki-tabu
Agr1-Pst-buy every 7-book
'she bought every book'

b. *kila ki-tabu, a-li-(ki)-nunua [t]
every 7-book Agr1-Pst-(Agr7)-buy [t]
'every book, she bought (it)'

Again, if the preverbal NP is a Topic as Zwart says, then a quantifier in this position should be ungrammatical. (51) clearly shows this is not the case.

(51) a. kila m-toto a-li-nunua ki-tabu
every 1-child Agr1-Pst-buy 7-book
'every child bought a book'

b. kila ki-tabu ki-li-nunuliwa na mtoto
every 7-book Agr7-Pst-be bought by a child
'every book was bought by a child'

§8.5 Idiomatic Expressions

Further evidence against a pronominal analysis – again pointed out by Keach – comes from the fact that idiom subjects (i.e. true subjects according to our operational definition) resist topicalization (Bresnan and Mchombo 1987). Here I give two more examples presented in Deen (2006) (his (9a) and (9b)).

(52) a. Ni-li-fikiri kuwa mtindi u-me-vaa Asha
Agr1-Pst-think that brew Agr3-Perf-cover A.
'I thought that Asha is drunk' (lit. I thought the brew has covered Asha)

b. *Mtindi ni-li-fikiri kuwa u-me-vaa Asha
brew Agr1-Pst-think that Agr3-Perf-cover A.
lit. '(as for) the brew, I thought that it has covered Asha'

As this idiomatic expression resists topicalization, we know that it is a true grammatical subject and not a Topic.

§8.6 Typological Evidence

One way in which clitic pronouns are distinguished from agreement affixes is that the clitic pronouns are free to move relative to the verb (or vice-versa). This can be seen in a variety of languages from Tagalog to languages in the Southern California branch of Uto-Aztecan, which have semi-unordered positioning of the verb and clitic.[14] In Swahili, however, the verbal complex acts as a single unit – the SM is entirely bound to the verb, always remaining in its

[14] Cf. Deen (2006, p. 231) for the cross-linguistic data.

original position, regardless of where the subject moves. This suggests that the SM is not a pronominal clitic. It is entirely consistent, however, to suggest that the SM is instead a copula in the auxiliary verb slot, an item which typologically are also not free to move relative to the main verb (e.g. the ungrammaticality of sentences like 'The dog going was.')

§9. Implications of a Pronominal Analysis: The Split-Verb Hypothesis

Zwart's resumptive pronoun analysis, when taken in conjunction with facts about the status of the OM in Swahili, requires a significant reanalysis of the verbal paradigm. There is a general consensus in the literature (Bresnan & Mchombo 1987; Allan 1983; Wald 1979) that the OM is also a resumptive pronoun, conditioned by certain discourse factors.[15]

This creates an interesting problem for the verbal paradigm, namely that under a combination of Zwart and Bresnan & Mchombo's analyses, the tense marker occurs between two resumptive pronouns. Zwart rightly suggests that this casts doubt on the tense marker as an affixal part of the verb structure, for it is typologically unusual for a tense marker to affix to a pronominal element. Only if the tense is marked on an auxiliary is it perhaps possible for a pronominal element to intercede between the tense morpheme and the main verb, such as in (53). Even here, the evidence is weak.

(53) Daniel was himself going.

So this is precisely the analysis Zwart would like to take – that Swahili verbs actually consist of an auxiliary and a main verb, separated at times by the pronominal clitic, and that the tense marker is in fact a full verb. His analysis is compared to the traditional one in (54).

(54) *Traditional Analysis*
SM + Tense + Relative + OM + Stem + Mood + Relative

Zwart's Analysis
Pronoun + Auxiliary + Relative Pronoun + Verb + Mood + Relative

Historically, there is decent evidence to support this. It is clear that many – if not all – the tense markers in Swahili arose through the phonetic reduction of full verbal forms. A list of some of the common forms and their derivation is given in (55).

(55) *Etymology of Swahili Tense Markers*

-na-	Present	Conjunction/Preposition *na* 'and/with'
-li-	Past	Copula *–li* 'be' (cf. McWhorter 1992)
-ta-	Future	*-taka* 'want'

And there are similar – though less certain – etymologies for the various other tenses, so that most are assumed to have been derived from full lexical verbs, if at varying points in the past.

[15] Conditions in which the OM is obligatory include: non-standard word-order for the object NP; when the object NP is the Topic (which is in turn related to displacement of the object NP); when the object is animate.

The above forms, however, are extremely well-motivated.[16] Zwart believes this is evidence for the claim that the tense markers are in fact "still verbs" which host the cliticized subject and object pronouns, and there is some reason to believe that Zwart may be onto something.

For starters, notice the striking parallels between the auxiliary and the main verb:

(56) Clitic + Auxiliary + Relative Clitic + Verb + Relative

Swahili has three ways of forming relatives: the first is via the independent complementizer *amba-* 'who/which/that', which takes a relative suffix. Alternatively, that same relative suffix can either be suffixed to the verb stem, or prefixed immediately following the tense marker. The examples in (57) illustrate these uses respectively.[17] Note that each of these relatives has a different use, but the differences need not concern us here.

(57) a. vi-tu amba-vyo Juma a-li-vi-fany-a
8-thing Comp-Rel8 J. Agr1-Pst-Agr8-do-Ind
'the things which Juma did'

b. vi-tu a-li-vyo-vi-fany-a Juma
8-thing Agr1-Pst-Rel8-Agr8-do-Ind J.
'the things which Juma did'

c. vi-tu a-vi-som-a-vyo Juma
8-thing Agr1-Agr8-read-Ind-Rel8 J.
'the things Juma does'

Zwart would analyze the relative in (57b) as attaching to an auxiliary (the past tense *-li-*), followed immediately by a word boundary.

Like in English, it appears that the 'auxiliary' verb triggers morphological changes on the main verb. When present, it triggers the indicative mood; when absent, the mood of the verb can be subjunctive, imperative, and sometimes even negative (although the negative occurs with the 'auxiliary' as well). Examples (58) and (59) compare this phenomenon in English and Swahili (following Zwart's analysis).

(58) *English*

a. he *was* kick-*ing*
Pronoun Auxiliary Verb-Participial

b. he Ø kick-*ed*
Pronoun (no auxiliary) Verb-Tense

[16] *Na* still exists in the language as the particle 'and' or 'with'; McWhorter (1992) gives an excellent account of the etymology of *li* and its origins as a copula; and it is generally known that typologically, future tense verbs/markers often arise from forms of 'want' or 'will' (cf. English, where 'to will' originally had the sense of 'to want', before becoming our modern future tense marker).

[17] Note that in all these cases the presence of the OM in addition to the overt object NP indicates a heavy emphasis (perhaps Topicalization) of the object, i.e. 'the *things*, the ones Juma does'.

(59) *Swahili*

a. a-na- -sem-a
Pronoun-Auxiliary Verb-Indicative
'he is speaking'

b. Ø sem-eni
(no auxiliary) Verb-Imperative
'speak! (pl.)'

c. Ø a-sem-e
(no auxiliary) Pronoun-Verb-Subjunctive
'he may speak'

Further evidence comes from the orthography, where verbs sometimes follow a disjunctive tradition. Thus one will see both (60a) and (60b) in free distribution in everyday writing.

(60) a. u-na furaha
Agr1-'with' happiness
'you are happy' (lit. 'you are with happiness' or 'you have happiness')

b. u-na-furaha
Agr1-Pres-be happy
'you are happy' (no second interpretation)

Finally, there is strong phonological evidence for the Split-Verb Hypothesis. Swahili has consistent penultimate stress, and this stress can determine word boundaries. When suffixes are added to words, this stress shifts to reflect it:

(61) a. **jí**ko 'kitchen' > ji**kó**ni 'in the kitchen'
b. **só**ma 'read! (sing.)' > so**mé**ni 'read! (pl.)'

Interestingly, the secondary stress is also subject to this phenomenon:

(62) a. m-tu **á**-li-enda sokoni
1-person Agr1-Pst-go market
'the person went to the market'

b. m-tu a-**lí**-ye-enda sokoni
1-person Agr1-Pst-Rel1-go market
'the person who went to the market'

This would seem to suggest a word boundary after the tense marker in (62a) and after the relative in (62b) – exactly as Zwart's hypothesis predicts.

While I find this evidence for a split verb incredibly convincing, Zwart's analysis of the tense markers seems patently wrong. No native speaker actually analyzes the marker this way.[18] The tense marker long ago lost the lexical meaning it once had. But if Zwart's resumptive pronoun hypothesis is correct, then we are forced to accept his analysis of the tense marker as well, which to me is a strong argument *against* a pronominal analysis. Why? Because I think it is fairly easy to understand where Zwart went wrong: he has located contemporary reflexes of what was originally a pronoun in the language and – sometimes quite convincingly – argued that this analysis still holds for modern Swahili.

Clearly, however, this analysis lost its validity some time ago. Zwart got the argument essentially right, but he applied it to the wrong era. I *do* think that Zwart has given us an excellent analysis of the Swahili verbal paradigm just prior to or in the beginning stages of Early Modern Swahili (as defined by McWhorter 1992) – that is, just before some of the earliest written records of Swahili. By the time of these records in the 17th century, the resumptive pronoun had just begun its incorporation into the verbal paradigm via its reanalysis as a copula. Today, both the resumptive pronoun and the separate auxiliary verb have been entirely grammaticalized onto the verb (a process which I will attempt to detail momentarily).

In the same way as Zwart, one can utilize leftover reflexes of any recently old phenomenon in language to support an analysis which includes that phenomenon; the analysis will simply be outdated. We can, for instance, argue that the SM is in fact a suppletive copular verb which shows agreement with a full subject NP, which would be the intermediate step between a pronominal analysis and an agreement analysis. Obviously this too requires a reanalysis of the verbal paradigm in a mode different from either Zwart's analysis or the standard accepted today. But we can find evidence for this in the language.

One might ask, 'if a copular analysis is outdated in the same way as a pronominal analysis, why suggest it?' There are two responses to this. First, it is of historical interest to understand the process by which the modern verbal paradigm developed in Swahili, and I believe a copular analysis provides a crucial stepping stone. Second and more importantly, I would like to suggest that this analysis is in fact of contemporary interest; that is, a copular analysis is the best way to analyze the verbal paradigm in *modern* Swahili. Why think this is the case? Some authors (Deen 2006; Bresnan & Mchombo 1987; Keach 1995) have argued for a point mentioned earlier – that the Swahili SM still exists as a hazy middle ground between pronoun and agreement marker (at least in some dialects of the language). That middle ground to me appears to be a copular analysis of the SM in Swahili.

§10. The Copular Analysis

I propose that we analyze the Swahili verb as follows:

(63) *The Copula Hypothesis*
Copula + Tense + Relative (+) Object Pronoun + Stem + Mood + Relative

The optional morpheme division (+) is meant to indicate the optionality of following the Split-Verb Hypothesis and positing a word boundary between the relative and the OM. In this case, we would analyze the initial prefixes as constituting one auxiliary verb 'be', followed by the main

[18] My own personal experience assures me of this. Having been aware of the origins of the *-ta-* tense marker, I mentioned this to a number of native speakers of Swahili, and all were surprised to know that it had once been a verb.

verb consisting of the OM, stem, mood, and relative. Essentially, this entails analyzing every Swahili verb using a paradigm something like 'be + participial'.

Unlike Zwart's version of the Split-Verb Hypothesis (where the tense marker is in fact a full verb), I believe that the copular analysis is in fact salient to speakers of Swahili, and that they do in fact conceptualize predication in this manner (i.e. as a description more than an event). Here I draw upon the tradition among many functionalist linguists of a continuum of possible lexical semantic meanings, with nouns at one end of the continuum and verbs at the other. It would seem that Swahili verbs fall in the range of 'verby adjectives'. My point here is that in many languages, categories like 'noun' and 'verb' are far from being clear-cut and distinct; this is also true of Swahili.

§10.1 Evidence for the Copular Analysis

There is some empirical support for this in the language. First, as mentioned at the start of this paper (cf. (5)), the various forms of the SM and those of the nominal prefix are similar in a number of cases. Six of the sixteen noun classes show the same morpheme for both subject agreement and nominal agreement. I argue that these two sets of morphemes in fact serve the same function, so that constructions like (64a) mirror that in (64b), and that the *a-* prefix is simply an alternate form of the *m-* prefix – or at the very least, that they share the same functional load.

(64) a. m-tu m-kubwa
1-person 1-big
'a big person'

b. m-tu a-na-enda
1-person 1-Pres-go
'the person is going'

This is in fact my weakened assertion: whatever the SM is, it is what 'derives' the verb, or *makes* the verb a verb, in the same way that the nominal prefix derives a noun and qualifies it as a noun. It is *not* the stem which determines its lexical status as a verb, but rather the verbal prefix (and here I am going to adopt this terminology to replace SM, so as to illustrate the parallels between the verbal prefix and the nominal prefix; I will also drop the use of Agr in my glosses, opting instead for just the number of the noun class). Thus it is sensible to conceive of the verbal prefix as *the* verbal element in the verb paradigm, and analyzing the verbal prefix as a copula fits this role (my stronger claim).

The verbal stem does not determine its lexical status as a verb because it may adopt the nominal prefix instead. Deverbatives (nominal derivation from verbs) in Swahili are accomplished by prefixing the nominal morphology to the verbal stem, often accompanied by a derivational suffix as well (although this does not occur in some contexts). Thus the difference between (65b) and (65c) – by the traditional analysis – is a conditioned by the choice of either nominal or verbal morphology, *not* by any special lexical status given to the verb root. Put differently, there is nothing about the verb root (aside from the action conveyed by its semantics) which makes it inherently verbal (or nominal).

(65) a. m-toto
1-small
'a small (person)'

b. m-ganga
1-heal
'a healing (person)' (trad. 'healer)

c. a-na-ganga
1-Pres-heal (tr.)
'a healing (person)' (trad. 'he is healing')

I of course would like to conflate this difference further, and claim that *both* the *a-* and *m-* prefixes are a type of predicative prefix. The difference between (65b) and (65c) is not a difference of verbal versus nominal morphology, but rather the addition of tense. Hence the glosses in (65): a 'small person' for (a), and a 'healing person' for both (b) and (c) ('person' because class 1 is animate). The structures are essentially the same.

There is evidence that the verb functions similarly to an adjective in certain non-standard uses of the relative. Relatives function essentially as modifiers, and especially so in Swahili where the entire relative typically consists of a relativized verb in the complement position of a noun, as in (66a) (Standard Swahili).

(66) a. ni-li-kutana m-tu a-na-ye-itwa Hamisi
1-Pst-meet 1-person 1-Pres-Rel1-be called H.
'I met a person who is called Hamisi'

b. ni-li-kutana m-tu a-na-itwa Hamisi
1-Pst-meet 1-person 1-Pres-be called H.
'I met a person called Hamisi'

However, in some non-standard dialects, one hears forms such as (66b), where the relative marker is omitted. Note that the verb *–itwa* can in this context only be a modifier – no other interpretation is possible except 'the person who is called *x*' or 'the person called *x*'. This means that the verb *–itwa* serves as a modifier in (66) *even though it lacks the relative marker*. This is evidence that Swahili verbs function mostly like modifiers, rather than 'full' verbs.

(67a-f) shows how the noun-verb distinction in Swahili becomes fuzzy at times.

(67) a. mtoto ana mpira 'the child has a ball'
b. mtoto ana baridi 'the child is cold' (lit. 'has cold')
c. mtoto ana furaha 'the child is happy' (lit. 'has happiness)
d. *mtoto anabaridi 'the child is cold'
e. mtoto anafuraha 'the child is happy'
f. mtoto anaenda 'the child is going'

I argue that (67f) is parallel to (67c) and even (67a). This is most salient for the present and past tenses, but harder to motivate for the other tenses, as we will see later.

§10.2 Historical Evidence

First some historical background. Some of the information here has been presented in an earlier section (§6), in which case it will be glossed over cursorily. However such background is important for understanding my argument.

Swahili once had a full copular verb *–li* 'be' which is attested in a number of sources. Remnants of the form appear in Ashton's (1944) *Swahili Grammar*, and it is still present in the relative form of 'be' in some dialects today (e.g. *uliye* 'you who are', p. 205). By the time of E(arly) M(odern) S(wahili) (as defined by McWhorter 1992), the present tense uses of this verb were restricted to locative constructions, such as (68a) and (68b) below (Knappert 1969, 8). Other types of linking predicates were formed by simply omitting the copula, as in (69) (this too is retained in modern dialects). In all other tenses, *–li* was used categorically, i.e. regardless of whether the construction was locative or descriptive. Thus there were two strategies for linking predicates in EMS – use of the copula *–li* or omission of the copula.

(68) a. u-li-we mbali haka
you-Cop-you far very
'you are very far'

b. tu-li-po
we-Cop-Loc
'we are here'

(69) Hamisi m-pishi
H. 1-cook
'Hamisi is a cook'

Swahili also once had a set of full pronouns – sometimes called the Free Subject Prefixes (FSPs), reprinted below. These became the same resumptive pronouns which left reflexes in the language for Zwart to find, and – according to my analysis – developed into a copula and were later incorporated into the verbal paradigm.

(70) *The Old Swahili Pronoun (FSPs)*

	Singular	Plural
1	ni	tu
2	u	m
3	yu	wa

By the time of EMS, these pronouns were coming to be replaced by the modern pronouns (cf. (42), §8.1), so that the older pronominal form became restricted to use with the copula *–li* (as in (68) above), and as a resumptive pronoun co-occurring with the new form of the personal pronoun as in (24), repeated here.

(24) wewe u m-kurugenzi
you you/Cop 1-director
'you are the director'

Prior to written record, however, this resumptive pronoun became reanalyzed as a copula, because it filled the copula slot in the grammar (McWhorter 1992, 22-23). So before the time of EMS, the FSPs had lost their status as a resumptive pronoun, diverging in its use.

As late as 1944, Ashton still listed the FSP as an acceptable form of the copula, even occurring with non-animate noun classes, suggesting that analogy had expanded the copular use of the FSPs to the other noun classes as well (see (71) below). Today, the copula with the form of the FSPs has been replaced in the present tense by the non-suppletive copula *ni* and in the other tenses by a form of the verb *kuwa* 'be'. At the time of Ashton's grammar, however, this change was only just beginning. She lists *ni* as being used only in the third person, while the FSPs are acceptable only in the first and second person. The fact that the FSPs-as-copula was replaced by the *ni*-as-copula suggests that the FSPs took on a different functional role (i.e. their incorporation into the verbal paradigm), allowing for the takeover of the copula by *ni*.

(71) a. ki-su ki wapi? (Ashton 1944, 20)
7-knife Cop7 where?
'where is the knife?

b. u m-pishi?
Cop1 1-cook?
'Are you a cook?'

By my analysis of the SM, I would like to say that the FSP-as-copula was incorporated into the verbal morphology, and retained some of its predicating function, causing a conceptual reanalysis of Swahili verbs to something like 'be + participial'.

It is not difficult to sketch the process by which this occurred. We can illustrate it very easily with the present tense of the verb 'have'. The verb 'have' is formed by combining a form of 'be' with the particle *na* 'and/with', as in (72).

(72) u-na ki-su?
Cop-with 7-knife
'do you have a knife?'

But are we doing as Zwart did and applying an outdated analysis? Why not instead simply say that *–na* has been reanalyzed as a verb meaning 'have' (in this particular construction only), while *u-* serves as a simple agreement prefix? It doesn't appear that this is the right approach, for the infinitive form of 'have' is in fact *kuwa na*, where *kuwa* is the infinitive of the present verb for 'be', and *na* is of course the particle meaning 'with/and'. So the verb 'have' in Swahili is expressed using a construction meaning 'to be with'.

The next step on the way to incorporation into the verbal paradigm was the use of this 'have' construction with abstract concepts like 'happiness', so that *una furaha* could mean either 'you are with happiness' or 'you have happiness'. At some point, however, this construction had to have been reanalyzed as the stative 'be happy', because today we see the more synthetic

construction *unafuraha* which can only be analyzed as a verbal 'be happy'. This provides further evidence to the fact that *u-* plays some sort of linking or predicative role, and should be interpreted as 'be' plus some type of stative or participial.

By analogy, we can analyze verbs like *u-na-enda* from the stem *–enda* 'go' as meaning something akin to 'you are (in the state of) going'. Speakers of the language, too, would have made this analogy, and the paradigm of 'be + participial' would have broadened in its use. And it was not just this which drove the analogy – the same paradigm developed with the past tense form *–li*. We have already outlined this process briefly in §6, but we will need more detail here.

Remember that EMS had several strategies for forming linking sentences: omission of the copula, use of the FSPs as a copula, or the full verb *–li* meaning 'be'. We already know how the FSP became cliticized onto the verbal structure. *–li*, on the other hand, was reanalyzed over time as a past tense affix through an ambiguity in the language. This left a functional gap in the language concerning copulas, to be filled later by the modern form *ni*.[19]

The reanalysis of *–li* occurred as follows: in EMS, the imperfect tense affix was *-a-*. The third person singular form of the cliticized/resumptive pronoun was also *a-*. This meant that the forms *a-li* 'he-Cop' and *a-a-li* 'he-Imp-Cop' were phonologically identical (the latter written simply as *ali-*). This imperfect tense could be used either as a linking verb or as a past continuous tense, just as in English ('he was a child' versus 'he was sitting'). The EMS equivalents of these are given in (73) (adapted from McWhorter 1992, p. 31). Note here the ambiguity of the *a-* morpheme.

(73) a. a-li mtoto
he/Imp-Cop child
'he was still a child'

b. a-li me-keti
he/Imp-Cop Pst-sit
'he was sitting'

All this was confused further by the fact that Swahili also has a *present* tense marked by the morpheme *-a-*. As such the form *a-li* was entirely ambiguous between either present or past tense, and so *–li* eventually dropped out of use as a full verb entirely, and became incorporated into the verbal paradigm as a marker of past tense, while other copula strategies were developed in its absence.

Notice the interesting similarities between forms like those in (73) and ones such as *una furaha*. Both constructions consist of a form of 'be' followed by a description or predication; or, alternatively, both consist of an auxiliary followed by a main verb. This explains our evidence for the Split-Verb Hypothesis – it comes from the reflexes of this former state of the language. The Split-Verb Hypothesis is still essentially right in that there is an internal structure to the verb – an auxiliary/verbal component and a main verb/participial component.

With two highly prevalent tenses in the language now following this paradigm, it is a simple matter of analogy to understand the various other tenses in a similar way. More to the point, it is only a matter of analogy for the remaining tense paradigms to develop. The end result of that original verbal paradigm of 'auxiliary/copula + verb/predicate' is that modern Swahili verbs are

[19] Even this form is being replaced in some dialects (namely the Kimvita dialect of Mombasa) with the locative copula *–ko*, a form which takes the SM and thus eradicates the ambiguity of *ni*.

much closer to the nominal end of the spectrum, analyzable as something like 'be + participial'. Even though the auxiliary has been incorporated onto the verb, its subtle effect on the verb's meaning remains.

§11. Final Considerations

Let us reiterate this process with one more example, a still-emerging phenomenon in the Kimvita dialect of Swahili spoken in urban Mombasa.

In SS, the verb *kwisha* 'to finish' often appears as an auxiliary:

(74) simba a-me-kwisha ku-uliwa (Mohammed 2001, 84)
lion Agr1-Perf-finish 13-be killed
'the lion is already killed'

Most frequently, it occurs with the perfect tense marker *-me-* as follows.[20]

(75) u-me-kwisha ku-ona?
1-Perf-finish 13-see
'have you seen?'

Interestingly, the specific form *umekwisha* has become so common that it has grammaticalized somewhat and become its own tense, known as the *-mesha-* tense (Hinnebusch and Mirza 1998, 117), usually interpreted as 'have already':[21]

(76) u-mesha-ona?
1-already-see
'you've already seen (it)?'

The above form made its way into the grammars before stabilizing, however, so that today expressions like (75) are rare. Instead the *-mesha-* tense has been phonetically reduced to just *-sha-*. At the same time, it has expanded its meaning to the metaphorical use of 'see' as in 'understand'. Hence:

(77) u-sha-ona?
1-already-see
'you have understood?/see?/understand?' (lit. 'you have already seen?')

This expression is a (ubiquitous) colloquialism, and I personally have never heard the use of *-sha-* outside the context of *-ona*, but it illustrates the point well. Here we see how the auxiliary verb grammaticalizes onto the main verb, creating a new tense/aspect affix in a relatively short historical span.

[20] The exact use of the *-me-* tense is controversial and blurs the lines between tense and aspect. At times it is even used to convey a sense of present (e.g. 'he has put on the shirt' is equivalent to 'he is wearing the shirt'). Fortunately, these tricky details need not concern us here; but see Mohammed (2001) for an excellent discussion of current uses.

[21] Note that *kwisha* is actually an irregular verb, made of a combination of the nominal prefix *ku-* (a noun class used to mark verbal infinitives and participials) and the stem *-isha*. This irregular use of the *ku-* prefix drops in (83).

§11.1 Some Objections

Of course, it is questionable whether these non-primary tenses (for our purposes, neither the present nor simple past tenses) are also analyzable in terms of the 'copula + participial' paradigm except via analogy to the primary tenses. Forms like *ushaona* and perhaps even the future *-ta-* (which derives from *-taka* 'want') don't lend themselves to the Copula Hypothesis, as they don't derive directly from copular forms like the present and past tenses do. Since *kwisha* in (83) here is being reanalyzed itself, we would not necessarily expect its internal morphology to be retained – although it might be. Why should we see (77) as anything except an example of grammaticalization, rather than the incorporation of a copula into the verbal paradigm? Why assume that the present and past tense paradigms were analogously expanded to the other tenses?

First, there is the simple fact that the SM on the verb derives from a copula and contributes to the analysis of *kwisha* as 'be + participial'. But as we just mentioned, the reanalysis of an auxiliary verb into a tense marker may include the reanalysis of the copula as well. Still, the evidence for a copular analysis in §10.1 lends credence to the idea that Swahili predication is not 'fully verbal', and this is easily explained by interpreting the SM as a verbal prefix, i.e. that which makes the stem a verb, functioning in much the same way as the nominal prefix which makes the stem a noun.

More evidence for this comes from the fact that verbs have their own special noun class, marked by the *ku-* prefix. This nominal prefix is used to mark infinitives and participials in the language, making the stems it affixes to at once both verbal and nominal – verbal in the sense that the semantics of the stem involve a state or action; nominal in that it functions syntactically as such and in that we interpret it as a participial. The proper analysis of (75) is not 'have you seen?' but rather 'you have finished seeing?', where *kuona* is a participial. So we have a verbal auxiliary followed by a nominal rather than verbal element. When this form cliticizes entirely, the verb *kwisha* loses its lexical meaning and becomes merely functional, a marker of aspect. (At least, this is what we must say if we wish to avoid Zwart's analysis of the tense marker as a full verb.) Analogous to the present and past tense paradigms, we should thus view the SM rather than the stem as the verbal element of the construction. This is clear from the fact that *-ona* is a stem only – requiring nominal or verbal prefixes to give it its function – and the fact that *-sha-* must be a grammatical marker of aspect, leaving only the SM to serve as the functionally verbal element of the word.

So while the Copula Hypothesis is strongest when seen through the present and simple past tenses, the lack of copular origins for the other tenses need not give us dire concern. It is still true that the SM derives itself – and fairly recently in the history of the language for that matter – from a copula, that Swahili verbs act in very similar ways and take similar morphology to Swahili nouns, and that the stem of a verb is not what makes it a verb (i.e. that the SM, rather than the stem, is the 'verbal' element of the word).

But perhaps critics of the Copula Hypothesis would argue an inference to the best explanation: They might say that it makes just as much sense to say that *-ona* is the 'truly verbal' element in the construction as it does to say that *u-* is a verbal copula which predicates the participial 'seeing'. For them, *-sha-* is simply a marker of aspect and *u-* is nothing more than an agreement marker. They would argue that an agreement analysis is now the overriding paradigm for Swahili speakers, and matches their conceptualizations, and that it has come to replace the 'be + participial' conceptualization. They claim that an agreement analysis – particularly of the non-primary tenses – has by analogy come to be the norm. They cite evidence for this analysis

(Deen 2006; Keach 1995), and point out that the SM is more frequently dropped in urban dialects of the language today, as it no longer bears any functional weight. The copular analysis, they would say, is outdated in the same way as Zwart's, only not quite to the same extent.

But to say this, and to overlook the pervasive empirical and conceptual evidence which shows that Swahili verbs function in remarkably similar ways to Swahili nouns, would be to disregard some of the uniquely distinctive conceptual differences between the Swahili language and our own. While I certainly agree that Swahili is headed towards a fully agreement-based verbal paradigm, it is not yet there. Additionally, there are countless dialects of the language which have all evolved differently and at different rates; this is to be expected from its status as a lingua franca. In the meantime, the Swahili language will continue to exemplify the fact that many of the lexical categories which seem so cross-linguistically pervasive to us in the English are in fact merely imposed. And while such an imposed analysis can be made to work, we lose the valuable potential insight that is to be gained from a reconsideration of how we think about the most basic categories in grammar.

§11.2 Implications for Nominal Morphology

If even the weak version of my hypothesis is correct – that the verbal prefixes and the nominal prefixes both serve the same essential function, assigning a nominal class to a stem – it would seem to have significant implications for how nominal morphology works in Swahili. To begin with, we would want to call the SM a type of nominal prefix, just as *m-* is the nominal prefix which assigns a stem to the animate class of nouns, or the *ku-* prefix is a nominal prefix which assigns a stem to the participial or verb-like class of nouns. We would say that *m-* (traditionally analyzed as a nominal prefix indicating animate class) and *a-* (traditionally analyzed as the SM for third person singular) are two forms of the same general morpheme. By now it should be clear what was meant when I stated earlier that *mtu mtoto* 'a small person' and *mtu anaenda* 'the person is going/the going person' are analogous constructions, where the latter element modifies the former.

If the strong version of my hypothesis holds – that the Copula Hypothesis is true and that we are still justified in analyzing the SM as a copula in modern Swahili – then the implications are of an entirely different class. The question becomes this: if we are right in analyzing the SM as a copula, and if the SM and the nominal prefixes serve the same function, then mustn't we also call the nominal prefixes a copula as well? This may not be as far-fetched as it appears. At the very least, it resolves the issue which Ashton raises (1944, p. 46) when she says that the 'adjectival' (what I would call nominal) agreement "virtually converts the 'Adjective' stem into a noun in apposition." Rather than dealing with the difficult syntax involved in positing two constituents at the same grammatical level, this changes the construction to involve a linking clause instead.

§12. Conclusion

What is the status of the SM in Swahili? As a set of morphemes, the SM has a complex history. I have shown here how an early resumptive pronoun became reanalyzed as a copula and then cliticized onto the Swahili verbal complex, leaving a 'grammaticalization chain' of hints in the language. Zwart picked up some of the earliest reflexes of this resumptive pronoun, and uses this evidence to argue convincingly that the SM must be more than just a marker of agreement. Yet in saying that the SM still holds status as a resumptive pronoun, Zwart clearly misses the mark. The analysis is simply impractical for modern Swahili.

Where Zwart went right, however, was in understanding that there is an internal structure to the Swahili verb, leading him to posit the Split-Verb Hypothesis. This hypothesis is correct insofar as it implies that the 'auxiliary' half of the verb must contain some verbal element; Zwart was simply wrong to say that this element was the tense marker. Instead, I have argued that the SM – having stemmed historically from a copula – is the underlyingly verbal element of the word. In addition to offering a practical alternative to Zwart's analysis of the tense markers, the Copula Hypothesis presented here also draws empirical support from the noun-like behavior of Swahili verbs, while at the same time incorporating Zwart's insights about the internal structure of Swahili verbs. Together these facts make feasible an analysis of the Swahili verbal complex which approximates the 'copula + participial' paradigm.

At the same time, it is clear that Swahili even now is moving away from this paradigm, towards a purely agreement-based one. Yet a weaker version of the Copula Hypothesis – that is, saying that the nominal prefixes and the verbal prefixes serve the same functional roles – still holds for dialects in which an agreement-based paradigm is the norm. This entails that it is not the word stem which gives the word its status as a verb or noun, but the derivational prefix. And these prefixes are inherently nominal, blurring the line between nouns and verbs in favor of the nouns. So ultimately the SM is best interpreted as that thing which derives a verby noun, which in turn modifies other nouns, and it is in this derivational sense of being what *makes* the verb that the SM is the truly verbal element of the Swahili verb paradigm.

Works Cited

Allan, K. 1983. *Anaphora, Cataphora, and Topic Focusing: Functions of the Object Prefix in Swahili.* Vol. 1, in *Current Approaches to African Linguistics,* edited by Ivan R. Dihoff, 322-564. Foris.

Ashton, E.O. 1944. *Swahili Grammar.* London: Longmans.

Bokamba, Eyamba G. 1981. *Aspects of Bantu Syntax.* Preliminary Edition. Urbana: Department of Linguistics, University of Illinois.

—. 1987. "Question Formation in Some Bantu Languages." Indiana University.

Bresnan, Joan, and L. Moshi. 1990. "Object Asymmetries in Comparative Bantu Syntax." *Linguistic Inquiry* (MIT Press) 21, no. 2: 147-185.

Bresnan, Joan, and Sam A. Mchombo. 1987. "Topic, Pronoun, and Agreement in Chichewa." *Language* 63, no. 4 (December): 741-782.

Chomsky, Noam. 1995. "Categories and Transformations." Chap. 4 in *The Minimalist Program,* 219-394. Cambridge: MIT Press.

Crystal, David. 2003. "Pronoun." *A Dictionary of Linguistics and Phonetics.* Malden, MA: Blackwell. 376.

Deen, Kamil Ud. 2006. "Subject Agreement in Nairobi Swahili." Edited by John Mugane, John P. Hutchinson and Dee A. Worman. *Selected Proceedings of the 35th Annual Conference on African Linguistics.* Somerville, MA: Cascadilla Proceedings Project. 225-233.

Demuth, Katherine, and Carolyn Harford. 1999. "Verb Raising and Subject Inversion in Comparative Bantu." *Journal of African Languages and Linguistics* 20, no. 1: 41-61.

Givón, Talmy. 1976. "Topic, Pronoun and Grammatical Agreement." In *Subject and Topic,* edited by Charles N. Li, 150-188. New York, New York: Academic Press.

Hinnebusch, Thomas J., and Sarah M. Mirza. 1998. *Kiswahili: Msingi wa Kusema Kusoma na Kuandika (Swahili: A Foundation for Speaking, Reading, and Writing).* 2nd. Lanham, Maryland: University Press of America.

Keach, Camillia. 1995. "Subject and Object Markers as Agreement and Pronoun Incorporation in Swahili." In *Theoretical Approaches to African Linguistics*, edited by A. Akinlabi, 109-116. Trenton, New Jersey: African World Press.

Knappert, Jan. 1969. "The Discovery of a Lost Swahili Manuscript from the Eighteenth Century." *African Language Studies* 10: 1-30.

Lasnik, Howard, and Tim Stowell. 1991. "Weakest Crossover." *Linguistic Inquiry* 22: 687-720.

Machobane, 'Malillo. 1987. "The Sesotho Locative Constructions." *McGill Working Papers in Linguistics.* Montréal: McGill University. 33-52.

McWhorter, John H. 1992. "Ni and the Copula System in Swahili: A Diachronic Approach." Edited by E.F. Konrad Koerner. *Diachronica* 9, no. 1 (Spring): 15-46.

Mohammed, M.A. 2001. *Modern Swahili Grammar.* Nairobi: East African Educational.

Perrott, Daisy Valerie. 1951. *Teach Yourself Swahili.* London: The English University Press.

Rizzi, Luigi. 1992. "Early Null Subjects and Root Null Subjects." In *Language Acquisition Studies in Generative Grammar*, edited by T. Hoekstra and B.D. Schwarz, 151-176. Amsterdam/Philadelphia: John Benjamins.

Vitale, Anthony J. 1981. *Swahili Syntax.* Foris: Dordrect.

Wald, B. 1979. *The Development of the Swahili Object Marker: A Study in the Interaction of Syntax and Discourse.* Vol. 12, in *Syntax and Semantics*, edited by Talmy Givón, 509-524. New York: Academic Press.

Zwart, Jan Wouter. 1997. "Rethinking Subject Agreement in Swahili." NWO/University of Groningen.