

From Focus Marker to Copula in Swahili

Author(s): John McWhorter

Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: Special Session on Historical Issues in African Linguistics (1994), pp. 57-66

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via <http://linguistics.berkeley.edu/bls/>.

The Annual Proceedings of the Berkeley Linguistics Society is published online via [eLanguage](#), the Linguistic Society of America's digital publishing platform.

From Focus Marker to Copula in Swahili

John McWhorter
University of California, Berkeley

1. Introduction

This paper addresses the diachronic development of the copula *ni* in Swahili. Although today this item links subject and predicate in all persons and most predication types, historical and synchronic evidence indicate that this situation has only gradually arisen over the past few centuries. As recently as a century ago, the usage of *ni* was more constrained; it shared the copular domain with various other strategies. Diachronic developments since then have created a copula system anomalous among the Eastern Bantu languages, and I will show that while certain aspects of this development follow established patterns of syntactic change, other aspects, as well as the rate of the aforementioned changes, resulted from the widespread use of Swahili as a second language.

2. The Copula in Modern Swahili

It is extremely common for languages to morphologically encode a distinction between two copular subdomains: the equative, as in the Ewe *É-nye fia* 'I am the chief', and the locative, as in the Ewe *Me-lè xo me* 'I am in the house' (Kofi Agawu, p.c.). Swahili makes such a distinction, in that the copula *ni* is used in the equative, while in nonverbal locative predications, the copula is expressed with zero:

- (1) Hamisi **ni** mpishi.
Hamisi COP cook
'Hamisi is a cook.'
- (2) Hamisi **yu-ko** nyumba-ni.
Hamisi AGR-LOC house-LOC
'Hamisi is in the house.'

Ni is used in all persons and numbers. What is interesting is that in the spoken language, the omission of *ni* is grammatical (3). Informants suggest that this is particularly common after independent pronouns (4) and determined NPs (5).

- (3) Hamisi \emptyset mpishi.
Hamisi cook
'Hamisi is a cook.'
- (4) Sisi \emptyset wa-pagazi.
we PL-porter
'We are porters.'
- (5) Mi-zigo hii \emptyset mi-zito.
PL-load these AGR-heavy
'These loads are heavy.'

Informants universally concur that this omission is considered substandard; it is exceedingly rare in written texts. However, the fact that this construction occurs is a vital clue that zero copula was the original construction, given that a change from state A to state B in a grammar is generally mediated by a stage where A and B coincide, in combination with the rarity of attestations of the disappearance of copulas in natural change.

3. The Copula in Early Modern Swahili

Indeed, in Swahili texts of the seventeenth and eighteenth centuries¹, the language of which I will henceforth designate Early Modern Swahili (EMS), equative sentences quite regularly have zero copula. It is especially significant that we encounter this in writing, which suggests that the construction had yet to acquire a sociolinguistically pejorative connotation:²

- (6) *Pili* \emptyset *Rahamani Muwawazi wema na wawi.*
 second-word merciful merciful good and bad
 ‘The second word is the merciful, merciful for the good and the bad.’
 (Knappert 1969:4)
- (7) *Moli wetu* \emptyset *Rabbi tangu insi* *na ajinani.*
 Lord our Lord since mankind and demons
 ‘Our Lord is the Lord of men and demons since time immemorial.’ (ibid. 14)

Meanwhile, in locative sentences, the verb *li* is used, in contrast to the modern (2):

- (8) *U-li-we* *mbali haka...*
 you-be-you far very
 ‘You are quite far removed...’ (Miehe 1979:216)
- (9) *Tu-li-po.*
 we-be-LOC
 ‘Here we are.’ (Knappert 1969:8)

In non-present tenses, *li* is used in equative sentences as well:

- (10) *A-a-li* *mtoto tifuli.*
 she-IMP-be child still
 ‘She was still a small child.’ (Miehe 1979:216)

We see a cognate configuration in modern ChiBemba, in which zero copula is still grammatical in many equative constructions, although there is a *ni* cognate as well, and in the meantime, *li* is used in locative predications:

ChiBemba:

- (11) a. *Ma-kòndé.*
 they-nets
 ‘They are nets.’ (Sadler 1964:467)

b. *Ishina liakue ni Càali.*
 name his COP Caali
 'His name is Caali.' (ibid. 155)

c. *Tù lì kuno.*
 we COP here
 'We are here.' (ibid. 156)

4. The Development of *Ni*

Evidence of all kinds suggests that the copula *ni* is derived from the presentative reflex of *ni* used as a third person deictic in subject position:

(12) *Ni mamoja kwangu.*
 it-is the-same to-me
 'It's all the same to me.'

This usage would have been reanalyzed as a copula as a result of usage as a resumptive in topic-comment constructions. Such constructions, when used in contrast to subject-predicate constructions, highlight the identity of the predicate to the topic:

(13) [*Vita*] [*ni*] \emptyset [*taabu*].
 topic subject predicate
 TOPIC COMMENT
 'War, that's trouble.' / 'What war is is trouble.'

Over time, the force of the highlighting would have eroded via usage. Concomitantly, the topic would come to be processed as a subject, a common tendency in language change, forcing a reanalysis of *ni* as a copula as it necessarily moved into the copula slot after vacating the subject position for the erstwhile topic. Thus such sentences are now processed as subject-predicate sentences:

(14) [*Vita*] [*ni*] [*taabu*].
 subject copula predicate
 'War is trouble.'

Support for this scenario comes from various quarters. First, this type of development has been richly documented cross-linguistically. For example, Archaic Chinese had zero copula sentences (15) which contrasted with topic-comment constructions using a proximal demonstrative (16):

Archaic Chinese:

(15) *Wáng-Tái \emptyset wù zhe ye.*
 Wang-Tai outstanding person DEC
 'Wang-Tai is an outstanding person.' (Li & Thompson 1977:421)

(16) *Jī yù qí sheng yòu yù qí si, shì huò ye.*
 already wish him live also wish him die this indecision DECL
 'Wishing him to live while wishing him to die, that is indecision.' (ibid. 424)

By the first century, *shì* appears in unambiguously copular usages, and this usage is regular today:

First century Chinese:

- (17) Yú **shì** suo jià fu-ren zhi fù ye.
I be NOM marry woman GEN father DECL
'I am the married woman's father.' (ibid. 426)

A similar process has been documented for languages as diverse as Hebrew (Li & Thompson 1977:427-431), Panare (Gildea 1993), and Saramaccan Creole (McWhorter 1993).

We would expect that between the stage where *ni* served to highlight identity and its unambiguous processing and the stage where *ni* came to be processed as a copula, sentences with *ni* would retain a degree of the highlighting semantics despite having advanced significantly towards copular processing. The grammar by Delaunay (1885) apparently captures this stage in describing *ni* as a copula serving to lend attention to the predicate (161):

'[To] answer this question: "Who is your brother?", if I say "My brother is a sultan", then it is the notion of the sultan that carries the weight of the sentence, and I say *Ndugu yangu ni sultani* [My brother is a sultan]. On the other hand, to answer this other question: "Is everybody ready?", if I say "Us, we are ready...", it is as if I said: "As to us, we are ready," and I will say: *Tu tayari* [We are ready]'

We find further support for the pathway described above in another non-copular synchronic reflex of *ni*. The presentative *ni* can be analyzed as one manifestation of an element of more general syntactic movement which can be characterized as a general focus marker, occurring before a constituent which the speaker desires to call particular attention to:

- (18) a. Momela Farm **ni** i-me-fany-wa National Park.
Momela Farm FOC AGR-PAST-make-PASS National Park
'It is Momela Farm that was made into a National Park.'
(Closs, Kondo & Mbaye 1967:24)

- b. Momela Farm i-me-fany-wa **ni** National Park.
Momela Farm AGR-PAST-make-PASS FOC National Park
'What Momela Farm was made into is a National Park.' (ibid.)

It is clear that this usage is not copular, in that it can occur not only with full verbs as above, but also with verbs 'to be':

- (19) Juma a-li-kuwa **ni** mkurugenzi.
Juma AGR-PAST-be FOC director
'Juma was the *real* director.' (Wesana-Chomi 1978)

However, the semantics of a sentence like (19) are significant in that we see *ni* again used in a deictically highlighting function. The difference between (19) and sentences like (13), which produced the copula *ni*, is simply that in (19) the full verb 'to be', required in a non-present sentence, blocks *ni* from reanalysis as a

copula. Thus (19) reflects the semantic contribution that *ni* presumably made to present-tense nonverbal sentences before its reanalysis as a copula:

(20) a. EMS:

[Vita] [ni] \emptyset [taabu].
 topic subject predicate
 TOPIC COMMENT
 'War, that's trouble.'

b. Juma a-li-kuwa **ni** mkurugenzi.
 Juma AGR-PAST-be FOC director
 'Juma was the *real* director.'

Moreover, the historical documentation in combination with grammars over the past one hundred years support the hypothesis in rather neatly indicating a progression from zero copula as regular to *ni* as regular. The earliest text extant, the Hamziya, has no occurrences of *ni* as a copula. Although this hardly confirms that *ni* had not begun the process of reanalysis, its absence here is nevertheless suggestive. In 18th century texts, copular *ni* alternates with zero, with the latter somewhat more frequent. As late as 1900, Seidel describes zero copula as the default condition, then proceeding to outline alternative strategies such as *ni*. By the mid-twentieth century, grammarians such as Ashton in 1944 and Perrott in 1951 describe zero copula as secondary to the usage of *ni* and restrict it to certain grammatical contexts; by the late twentieth century grammarians describe it as marginal if at all.

A final piece of evidence supporting my claim is that according to the scenario I have sketched, *ni* would have been first reanalyzed as a copula in third-person sentences, given the third-person domain of the presentative *ni* used as in (12). This is confirmed by Ashton, who describes *ni* in 1944 as used only in the third person. Equally important is that cognates of *ni* in related languages are typically restricted to third person, as we see in ChiBemba and Kikuyu:

ChiBemba:

- (21) a. Ishìna liakue **ni** Càali.
 name his COP Caali
 'His name is Caali.' (Sadler 1964:155)
- b. Tù lì bàana be sùkuulù.
 we COP children of school
 'We are students.' (ibid. 156)

Kikuyu:

- (22) a. Muti **ni** mwega.
 tree COP good
 'The tree is good.' (McGregor 1905:121)
- b. N-**de** mwega.
 I-COP good
 'I am good.' (ibid.)

Thus a rich array of evidence demonstrates that the copular *ni* emerged as a result of its usage as a resumptive subject in topic-comment constructions.

5. The Evolution of the Locative Construction

On the other hand, locative constructions underwent a completely separate restructuring at the same time. As we have seen, the locative copula in EMS was the verb *li*:

- (23) *N-a-li-mo usindizi.*
I-IMP-be-LOC sleep
'I was asleep.' (Miehe 1979:216)

Today, however, *li* serves a past tense marker:

- (24) *Ni-li-nunua kitabu.*
I-PAST-buy book
'I bought a book.'

The reanalysis of *li* as a past tense marker most likely occurred within the context of the EMS imperfect, which was encoded with the infix *-a-*:

- (25) *A-a-li mtoto tifuli.*
she-IMP-be child still
'She was still a small child.' (Miehe 1979:216)

The reanalysis was spurred by the co-existence of two phenomena. First of all, the third person pronominal was (and is) *a*, creating in the spoken language a potentially ambiguous portmanteau morpheme of frequent occurrence in third person imperfect sentences as in (25). Concomitantly, it would appear that the verb *li* was also regularly used in imperfect constructions, a situation made likely by the fact that *li* is regularly used in this fashion in sister languages such as ChiBemba:

ChiBemba:

- (26) *Tù à lí mona màayo.*
we PAST be see mother-my
'We saw my mother.' (Sadler 1964:55)

The result would have been that in the EMS equivalent of (26), shown in (27), the stage would have been set for the reanalysis of *li* as a past marker, with new generations processing the initial *a* as simply a pronominal, with no imperfect semantics.

Reconstructed EMS sentence:

- (27) *A-a-li ona mtu.*
he-IMP-be see man
'He saw a man.'

When *li* was reanalyzed as a past marker, its semantic contribution to present-tense locative predications would have become opaque, and to maximize transparency, it dropped out of such constructions. Thus emerged the modern locative construction as in (2), which had already been regularized by the late nineteenth century, as we see in (28):

Late nineteenth century:

- (28) Sasa **tu-ko** katika nchi nyingine.
 now we-LOC in country different
 'Now we were in another country.' (Harries 1965:124)

It is highly likely that it was indeed the reanalysis of *li* as a past marker which spurred the disappearance of *li* from locative constructions, given that *li* retains its locative function in the sister languages which have not incorporated *li* into the tense-affix paradigm, as we see in sentences such as the ChiBemba (11c).

Thus we see how two independent processes have neatly reversed the division of labor in the Swahili copula scenario: where in EMS, equative predication was expressed with zero while an overt element was used with locatives, today an overt element is regularly used in equative predications while zero is the rule with locative ones.

EMS:

Hamisi \emptyset *mpishi*.
 Hamisi cook
 'Hamisi is a cook.'

Yeye *yu-li-ko* *nyumba-ni*.
 he AGR-COP-LOC house-LOC
 'He is in the house.'

MODERN:

Hamisi **ni** *mpishi*.
 Hamisi COP cook
 'Hamisi is a cook.'

Yeye **yu-ko** *nyumba-ni*.
 he AGR-LOC house-LOC
 'He is in the house.'

6. Swahili Diachrony and Second-Language Acquisition

Standard Swahili is generally spoken only in official contexts, except by a minority of coastal Muslims. It is a second language for most of its speakers, serving as the primary vehicle of interethnic contact in East Africa. I will finally discuss how the phenomena described above are manifestations not only of explainable tendencies in syntactic and semantic change, but also of the effects of second-language acquisition upon the Swahili grammar.

In the process of its evolution, *ni* has succeeded in marginalizing what was a competing copular strategy in earlier stages of the grammar. In EMS and up until about the turn of the century, free-standing subject prefixes, reflecting the earlier status of those prefixes as independent pronouns, could be used to associate subject and predicate as well:

- (29) Hamisi **yu** *mpishi*.
 Hamisi he cook
 'Hamisi is a cook.'

- (30) *Uwe* *kifungo-ni* *u* \emptyset *kijuwo muutasima*.
 you knot-LOC you book precious
 'Thou art in that unison the well-guarded book.' (Knappert 1968:64)

This is another common source of copulas cross-linguistically, currently occurring, for example, in Hebrew:

Hebrew:

- (31) David **hu** ha-ganav.
David he/COP the-thief
'David is the thief.' (Li & Thompson 1977:429)

Grammars throughout the century describe this construction as increasingly marginal; in 1967 Closs, Kondo & Mbaye found that out of twelve informants two rejected it altogether while the other ten found it archaic or substandard. By this time, the construction had apparently narrowed in its semantic scope to indicating particularly strong association between subject and predicate, but even here, it shared that domain with the dominant strategy, *ni*, as we see in (32).

- (32) a. Wewe **u** mtoto mbaya.
you you bad child
'You are a bad child (by nature).'
- b. Wewe **ni** mtoto mbaya.
you COP child bad
'You are a bad child (today or by nature).'

Given that equivalents of this strategy are more vigorous in many of Swahili's sister languages, it is possible that the swift and unequivocal generalization of *ni* was a function of it being more easily manipulated by second language speakers with only partial command of the morphology, given that the subject prefix strategy required command of six different prefixes. In support of this account, similar generalization of single copula morphemes is typical of reduced and restructured registers of Swahili. For example, urban Kenyans have generalized the locative construction we saw in (2) as a stative marker, such that it is used even with descriptive predicates:

- (33) Vipindi vya televisheni **vi-ko** boring.
shows of television AGR-LOC boring
'The television shows are boring.' (Myers-Scotton 1984:8)

Similarly, while the construction in (33) retains allomorphic variation, there are pidginized registers of Swahili in which a generalized item *iko* has been similarly overgeneralized but has even lost all vestiges of agreement, invariant across the board:

- (34) Yeye **iko** tayari.
he COP ready
'He is ready.' (Duran 1979:148)

Heine (1973:95-6) has described a register where *iko* is part of a new copula paradigm in which *ndiyo* is particularly emphatic, *ni* is neutral, and *iko* falls in between. A final indicator that second-language acquisition is a significant factor in the changes I have outlined is that in more isolated Northern dialects of Swahili, *li* is still used as a verb 'to be', in contrast to its incorporation as a past marker in the standard (Wald 1987:1005).

7. Conclusion

I have intended to demonstrate that the current copula configuration in Swahili is a recent development, representing a significant departure from the configuration more typical of languages of its family. The differences that Swahili displays are indeed due in part to established patterns of syntactic change. The development of resumptive deictic subjects in topic-comment constructions into copulas is well-attested cross-linguistically, while the incorporation of the verb *li* into the verbal affix paradigm is a predictable outcome of its usage in an imperfect construction which already contained an ambiguous portmanteau morpheme encoding both the third person singular and imperfect aspect. However, the speed and extent of these changes, which have left Swahili an anomaly amidst its more conservative sisters, are manifestations of the effects of centuries of second-language acquisition upon the grammar.

Notes

1. Sources: seventeenth-eighteenth century - Poetic Documents: *the Hamziya* (seventeenth), *the Kasida* (early eighteenth), *the Herekali* (late eighteenth) (as edited and translated by Knappert 1967, 1968, 1969 and Mische 1979); nineteenth century: "Safari Yangu ya Urusi na ya Siberia" (1896), by Salim bin Abakari (Harries 1965).
2. Early Modern Swahili citations will be given in italics.

References

- Ashton, Ethel O. 1944. Swahili grammar (including intonation). London: Longman.
- Closs [Traugott], Elizabeth, Ali Kondo, and Salma Mbaye. 1967. Some copula patterns in Standard Swahili. *Swahili* 37:5-25.
- Delaunay, (Le Père). 1885. *Grammaire Kiswahili*. Paris: F. Levé.
- Duran, James J. 1979. Non-standard forms in Swahili in West-Central Kenya. *Readings in Creole studies*, ed. by Ian F. Hancock, 129-51. Ghent: E. Story-Scientia.
- Gildea, Spike. 1993. The development of tense markers from demonstrative pronouns in Panare (Cariban). *Studies in Language* 17:53-74.
- Harries, Lyndon. 1965. *Swahili prose texts*. London: Oxford University Press.
- Heine, Bernd. 1973. *Pidgin-Sprachen im Bantu-Bereich*. Berlin: Dietrich Reimer.
- Knappert, Jan. 1967. *Traditional Swahili poetry*. Leiden: E.J. Brill.
- _____. 1968. The Hamziya deciphered. *African Language Studies* 9:52-81.
- _____. 1969. The discovery of a lost Swahili manuscript from the eighteenth century. *African Language Studies* 10:1-30.
- Li, Charles N. and Sandra A. Thompson. 1977. A mechanism for the development of copula morphemes. *Mechanisms of syntactic change*, ed. by Charles N. Li, 419-44. Austin: University of Texas Press.
- McGregor, A. Wallace. 1905. *A grammar of the Kikuyu language*. London: Richard Clay & Sons, Ltd.
- McWhorter, John H. 1993. *Towards a new model of genesis: Competing processes in the birth of Saramaccan Creole*. PhD Dissertation, Stanford University.

- Miehe, Gudrun. 1979. *Die Sprache der Älteren Swahili-Dichtung*. Berlin: Dietrich Reimer.
- Myers-Scotton, Carol. 1984. *Roots in Swahili? A locative copula becomes a stative marker*. Michigan State University Ms.
- Perrott, Daisy Valerie. 1951. *Teach yourself Swahili*. London: The English University Press.
- Sadler, Wesley. 1964. *Untangled CiBemba*. Kitwe, Northern Rhodesia: The United Church of Central Africa in Rhodesia.
- Seidel, August. 1900. *Swahili Konversations-Grammatik*. Heidelberg: Julius Groos.
- Wald, Benji. 1987. *Swahili and the Bantu languages*. The world's major languages, ed. by Bernard Comrie, 991-1014. New York: Oxford University Press.
- Wesana-Chomi, E. 1978. *A note on the apparent problem in the description of copula constructions in Kiswahili*. *Kiswahili* 48:38-40.