
**The discourse function of object marking in
Swahili and Makhuwa narratives**

Teresa Poeta

Proceedings of Conference on

Language Documentation and Linguistic Theory 4

Edited by Aicha Belkadi, Kakia Chatsiou and Kirsty Rowan

7-8 December 2013, School of Oriental and African Studies, University
of London

Hans Rausing Endangered Languages Project
Department of Linguistics
School of Oriental and African Studies
Thornhaugh Street, Russell Square
London, WC1H 0XG
United Kingdom

Department of Linguistics:
Tel: +44-20-7898-4640
Fax: +44-20-7898-4679
linguistics@soas.ac.uk
<http://www.soas.ac.uk/academics/departments/linguistics>

Hans Rausing Endangered Languages Project:
Tel: +44-20-7898-4578
Fax: +44-20-7898-4349
elap@soas.ac.uk
<http://www.hrelp.org>

© Teresa Poeta 2014

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, on any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the author(s) of that part of the publication, except as permitted by UK copyright law.

This publication can be cited as:

Teresa Poeta. 2014. The discourse function of object marking in Swahili and Makhuwa narratives. In, Aicha Belkadi, Kakia Chatsiou and Kirsty Rowan (eds.). *Proceedings of Conference on Language Documentation and Linguistic Theory 4*. London: SOAS. www.hrelp.org/eprints/ldlt4_09.pdf

The discourse function of object marking in Swahili and Makhuwa narratives

TERESA POETA

School of Oriental and African Studies

1. INTRODUCTION

Object marking in Bantu languages has been studied in depth both in terms of its morphological description and syntactic analysis (Riedel 2009; Marten and Kula 2012 *inter alia*). However, very little has been said on how it relates to discourse. This is in spite of common remarks which state a clear connection to topicality throughout a text (Bresnan and Mchombo 1987) and to the surrounding discourse in general. ‘... When the object noun phrase is omitted in a sentence the object marker will represent the missing object retrievable from previous discourse or context’. (Bearth 2003: 123). This study therefore looks at two Bantu languages: Makhuwa and Swahili, and explores the relationship between typological morpho-syntactic differences in object marking in the two languages and their discourse strategies. This paper examines how the two phenomena influence each other, as well as how discourse can contribute to our understanding of object marking more generally. To this aim a pilot study was conducted based on retellings of the Pear Story in Swahili and Makhuwa. Results and preliminary findings from this study are presented here.

2. LANGUAGE BACKGROUND

Swahili is spoken throughout East Africa as a lingua franca with a total of around 49 million second language users. Around five million speak it as their first language, mostly in the coastal regions of Kenya and Tanzania. **Makhuwa** is a language of northern Mozambique spoken by approximately 3,300,000 people across several provinces. Makhuwa-Meeto is one of eight main recognized Makhuwa varieties spoken in the Cabo Delgado coastal province by approximately 970,000 people (Lewis et al. 2014).

The basic grammatical structure of Swahili and Makhuwa and of Bantu languages more generally is characterised by SVO word order, a complex noun class system encompassing up to 18 noun-classes and a highly agglutinative morphology (Nurse et al. 2003). The verb form, most relevant for the following discussion, has a fairly fixed morphological template with slots occupied by subject and object markers, TMA, verbal roots and valency changing processes such as passive and applicative. In the Swahili example (1), the object *watoto* ‘children’ is co-indexed on the verb by the corresponding noun class object marker (OM) *-wa-¹*.

¹The abbreviations used in this paper are 1-2 SG/PL = first-second person singular/plural, 1-15 = respective noun class, CAUS = causative, DEM = demonstrative, FUT = future, NARR = narrative,

- (1) *Mwalimu a-ta-wa-som-esh* *watoto* *kiarabu*
 1.teacher 1.S-FUT-2.O-learn-CAUS 2.children 7.arabic
 ‘The teacher will teach children Arabic.’

3. OBJECT MARKING

3.1 Swahili

Swahili has a complex paradigm of object marking with a different object marker corresponding to each noun class (typical of Bantu languages). The object marker can co-occur with the full noun phrase but a comprehensive study of the precise circumstances under which this occurs has not yet been published. It has been often noted that proper names and animates are object-marked more frequently than others (Seidl & Dimitriadis 1997). More generally, the (non-) occurrence of an object marker is often optional and depends on a number of factors such as animacy or definiteness (Duranti 1979). Although object drop is not considered common for Swahili (Marten 2013), Seidl and Dimitriadis (1997) note that an object (including OM) can be omitted in spoken Swahili under certain circumstances.

3.2 Makhuwa

In Makhuwa object markers exist only for first and second person, and for classes 1 and 2. In these cases they are always obligatory, regardless of their meaning (Riedel 2009). No noun of any other noun class can trigger object marking on the verb.

- (2) (a) *ki-ni-m-wéha* *Hamisi /* *namarokoló /* *nancoólo.*
 1sg.s-pres.cj-1.o-look 1.Hamisi / 1.hare / 1.fish.hook
 ‘I see Hamisi/ the hare/ the fish hook.’
- (b) **ki-ni-m-wéha* *nveló /* *mokhorá /* *etthepó*
 1sg.s-pres.cj-(1).o-look 3.broom / 4.doors / 9.elephant
 (‘I see the broom/doors /carpenter/elephant.’)
 (Van der Wal 2009: 84-85)

This highly restricted object marking paradigm is unusual for Bantu languages and its effects on other areas of grammar remain poorly understood. Object drop, unlike in Swahili, seems to be quite common in Makhuwa as objects can be omitted if the speaker assumes it is clear from the context (Van der Wal 2010).

O = object, POSS = possessive, PRES = present tense, PRF = perfect, PRO = pronoun, PST = past, S = subject

3.3 Object marking paradigms and discourse

The connection of object marking to discourse strategies in Bantu languages has been noted, although it remains understudied. Bearth (2003) states that OMs can represent the missing object retrievable from context and Bresnan and Mchombo (1987) note the use of OM to maintain topics throughout discourse. Below is a Swahili example elicited with the MPI Staged Events stimuli (see section 4) where the speaker introduces the ‘bicycle’ with a full NP and subsequently refers back to it by using the appropriate OM². This is therefore a clear example of OM used for referent tracking in the given discourse.

- (3) *Kuna mwanamke ana baisikeli ya kukunja.*
‘There is a woman she has a foldable bicycle.’

Baisikeli y-ake, a-me-i-kunja,
9.bicycle 9.POSS.1 1.S-PRF-9.O-fold
‘Her bicycle, she folds it,’

imekuwa ndogo kabisa. Baadaye anakuja
‘it becomes very small. Later she comes,’

a-na-i-kunjua.
1.S-PRES-9.O-unfold
‘she unfolds it.’
[Swahili]

As can be seen from the previous sections (3.1 and 3.2), Swahili and Makhuwa differ greatly as far as their object marking paradigms are concerned. If one assumes the link of object marking to discourse exists, this typological difference between the two languages poses an important question. Does the presence/absence of object markers influence the discourse strategies of the language? If so, what does this tell us about our (so far mostly morpho-syntactic) understanding of object marking in Bantu languages?

4. DATA COLLECTION

In order to obtain comparable texts in the two languages new data was collected through primary fieldwork. This was especially needed as one of the two languages, Makhuwa-Meeto, is heavily under-documented and barely any material is available for the study of discourse and related phenomena. Two main techniques were used when collecting data for this study: the retelling of folk stories and the use of controlled stimulus material: Max Planck Staged events (Van Staden et al. 2001) and The Pear Story (Chafe 1980). This resulted in a comparable body of texts in the

² Only the relevant parts of the text are glossed. The rest is given as discourse context for greater clarity.

two languages. Furthermore, Swahili and Makhuwa speaking communities share a similar social and cultural background and there are no obvious major systematic differences in narrative conventions and storytelling traditions (in the sense noted by Bickel 2003 for the comparison of Himalayan languages).

The data used in this particular pilot study was elicited using the above mentioned controlled stimulus ‘The Pear Story’ and ‘Staged events’. These videos were shown to native speakers who subsequently narrated their content to another native speaker who in turn had no access to the video.

5. DATA ANALYSIS

5.1 Referential Density

Bickel (2003) defines Referential Density (RD) as the ratio of overt to possible argument NPs. This value is therefore calculated by identifying all available argument positions in the clause and comparing them with the amount of overt NPs found. There is considerable variation in the RD ratio across languages and the question arises of what this variation can depend on. An instance of a very low RD value would be a sample of language with a high occurrence of zero anaphora and frequent dropping of arguments.

Considering Swahili and Makhuwa-Meeto in this light helps us draw an initial picture of some of the differences in the ways the two languages refer back to participants in discourse. Here participants are all arguments of verbs. The ratio of overtly expressed arguments (i.e. by full NP, pronoun or OM) to all possible arguments of verbs is calculated giving us a percentage value of the RD for each text.

In terms of available tools to identify referents some preliminary observations can be made based on morpho-syntactic properties of the two languages. Noteworthy is the severely reduced object marking paradigm in Makhuwa as well as its relatively free object drop, in contrast to Swahili where this seems to be very restricted if possible at all. These properties should have a direct impact on the number of overtly expressed arguments in a text and might result in a lower RD. Results of the RD values of the text used in this pilot study are presented in Table 1.

Table 1
Referential Density results for The Pear Story

Language	Text	Possible argument positions	Overtly expressed arguments	RD
Swahili	<i>Pear Story</i>	91	83	91.2%
Makhuwa	<i>Pear Story</i>	82	57	69.5%

The results of RD values confirm the expectations that Makhuwa displays a lower referential density in its texts. A preliminary hypothesis is therefore that in Makhuwa as only a very limited group of nouns can be object marked and objects can be freely omitted, less arguments are expressed overtly resulting in a low RD overall.

5.2 Activation status of referents

The next step in relating object marking to discourse is to examine in more detail the activation status of referents in object position in these texts. This pilot study uses Chafe's (1994) activation status framework based on the cognitive efforts involved in the activation of a referent at a certain point in the narrative. Within this system a referent is considered GIVEN if it was already active, i.e. mentioned in the immediate preceding context and therefore needs the least effort to be activated. An ACCESSIBLE referent is one that has been mentioned already at some point in the past narrative and is therefore semi-active in the hearer's mind. Such a referent requires more effort to be reactivated than a given one. Lastly, a NEW, i.e. newly introduced referent to the text, is unknown to the hearer and thus inactive and will have to be activated implying the most cognitive efforts on the hearer's side. Such an examination helps to establish what kind of referents get object marked and which ones are referred to using different discourse strategies.

The aim of the activation status analysis is also an attempt to relate the use of OMs in discourse to Givon's (1983) hypothesis on the phonological coding of topics. Givon (1983) formulated a phonological coding weight scale where the more accessible a topic is the less phonological weight will be involved in its coding. See Table 2 below.

Table 2
Givon's (1983) phonological weight scale

<i>more accessible topics [light]</i>	<i>zero anaphora</i>
↑	<i>unstressed/bound pr. ('agreement')</i>
↓	<i>stressed/independent pronouns</i>
<i>less accessible topics [heavy]</i>	<i>full NPs</i>

5.3 Emerging patterns

In this section the most relevant examples are presented. The data is organised in pairs of comparable sentences in the two languages which show some of the common discourse patterns.

5.3.1. Object marking used for topic continuity

This first set of sentences exemplifies how object marking can indeed be used as a discourse tool in referent tracking and thus maintaining the topic continuity in both languages. This is of course possible only if an OM is available for that specific noun class. Since in Makhuwa only class 1/2 has corresponding OMs, sentences with referents from these classes were chosen. For the sake of comparison the same was done for the Swahili ones.

- (4) (a) **Swahili** *Ghafla anatokea mwanamume, anakupita karibu yake.*
‘Suddenly a man appears and passes by her.’

A-na-m-gonga,
1.S-PRES-1.O-bump
‘He bumps into her,’

sahani inaanguka chini na inavunjika vipande vipande
‘the plate falls down and it breaks into pieces.’

- (b) **Makhuwa** *Hovira. Ho-n-kuttumula,*
passed 1.S.PRF-1.O-bump into
‘He passed by. He bumped into her,’

hohiya nkhuki khuphwayaale
‘she let the plate, it didn’t break.’

In both cases the sentences are taken from the material elicited using the Staged events videos. In this particular video a woman stands with a plate in her hands. Suddenly a man appears and passing by bumps into her which results in her dropping the plate (which in one version breaks and in the other doesn’t). The woman is always introduced by full NP at the very beginning, however in the extracted sentences we can see that later on when she is mentioned again (in the object position) an OM is used to refer to her. This therefore exemplifies how an OM can be used to refer back to participants from class 1 and 2 already mentioned in the text and shows the use of Oms as discourse tools.

5.3. 2. Dropping of given objects

In the following set of sentences we can see how an object can be dropped in both languages. In both examples the object in question is mentioned in the preceding clause and is therefore very active. The speaker then simply drops it rather than referring back to it with a full NP, pronoun or OM (where available). This is not so surprising in the case of Makhuwa, where object drop has been noted to be common. However, in Swahili leaving out an object is not considered a frequent discourse strategy. This therefore shows that object drop in Swahili might be more common than assumed and that the two languages might vary less in this respect than was previously thought.

(5) (a) **Swahili** *Anachukua vidonge viwili vya sukari,*
'She takes two lumps of sugar,'

a-na-tumbukiza katika glasi.
1.S-PRS-put inside into glass
'she puts (them) into the glass.'

(b) **Makhuwa** *Ahocisa ikhwiyeri,*
'She took a spoon,'

a-ho-ttikhela mumkhu-ni
1.S-PRF-put plate-LOC
'she puts (it) on the plate.'

5.3.3. Topic continuity with object of non-1/2 noun class

The last set of data focuses on how discourse strategies co-occur with the typological differences among the two languages. The presented sentences involve objects belonging to noun classes other than 1/2 as it is here that Swahili and Makhuwa OM paradigms differ. In the Swahili example (6a), the object of interest is *kofia* 'hat' belonging to class 9. After first being mentioned with a full NP, this object is referred to by an OM in the subsequent section of discourse and is even dropped in the last clause³.

In example (6b) the exact same episode is described in Makhuwa. Similarly as in Swahili the hat (*ekofiyo*) is first introduced by a full NP. However, as the paragraph goes on, the hat continues to be referred to by a full NP. Considering what we have established about object drop in Makhuwa one might expect a given referent to be dropped as it is a continuing topic and would be easily understood from context. In contrast its numerous repetitions raise the question of why a full NP is chosen. Furthermore, this pattern in the data questions Givon's scale of phonological coding as presented in section 5.2 (Table 2). In this case *ekofiyo* 'hat'

³ In this case the OM present on the verb refers to the animate indirect object 'he' and as a consequence an OM for 'hat' would not be possible as Swahili allows only one object marker per verb

is a very accessible topic and the prediction is that it would be coded by a phonologically light form, contrary to what we find in our examples. A further question arises then of how this is linked to the non-availability of an OM.

(6) (a) **Swahili**

Akiwa mbele kidogo
‘When they were a bit ahead’

<i>wale</i>	<i>vijana</i>	<i>wa-na-i-ona</i>	<i>kofia</i>	<i>yake</i>
2.DEM	8.youth	2.S-PRS-9.O-see	9.hat	9.POSS.1

‘those boys see his hat’

Imeanguka chini. Kwa hiyo
‘it has fallen down. Therefore’

<i>wa-na-i-okota,</i>	<i>wa-na-m-rejishea</i>	<i>yeye</i>	<i>tena.</i>
2.S-PRS-9.O-pick up	2.S-PRS-1.O- return	3SG.PRO	again

‘they pick it up, they return (it) to him again.’

(b) **Makhuwa**

Tootho axilopwana phoonaaya wiira
‘And also the men saw that’

<i>ho-liyala</i>	<i>ekofiyo</i>	<i>awe.</i>	<i>Ph-aa-cisaya</i>	<i>ekofiyo</i>
1.S.PRF-forget	9.hat	9.POSS.1	NARR-PST-pick up	9.hat

‘he forgot his hat.They picked up the hat,’

<i>ph-aa-rwaaya</i>	<i>n-mahaa</i>	<i>uwara</i>	<i>ekofiyo</i>	<i>awe.</i>
NARR-PST-go	1.O-give	INF.wear	9.hat	9.POSS.1

‘they went to give(it) to him to wear his hat.’

<i>waa-rwaaya</i>	<i>n-maha</i>	<i>ekofiyo-awe, ...</i>
2.S.PST-PST-go	1.O-give	9.hat-9.POSS.1

‘When they went to give him the hat ...’

5.4 Preliminary analysis

From the data in this pilot study it appears that Makhuwa and Swahili can both use OMs as devices to track referents in discourse. This is true especially for class 1/2 as Swahili tends to object mark animate objects more frequently and for Makhuwa these are the only two noun classes where object marking is available.

It is in cases where the object nouns belong to other classes that the situation becomes more complex. In example (5) it appears that a given referent can be dropped in the immediate following discourse in both languages as it is still active in the hearer’s mind. This is perhaps more surprising for Swahili where object drop

is less frequent. However, it is the sentences in example (6) which shed new light on this discussion. In this case we have two different discourse strategies used in the same environment. In Swahili a given and highly topical object is introduced with the lexical NP and subsequently referred to by its OM. But in Makhuwa we have that same referent repeatedly coded by the full NP despite its high topicality and active status.

To summarise, it appears that in certain environments in Makhuwa there is only a binary choice of available tools for reference tracking, either to drop the object or to use the full NP. Moreover, repetition of lexical NPs seems slightly more frequent in the studied texts. This seems to occur consistently when in the corresponding Swahili environment an OM is used. On the other hand, when the object is omitted in Makhuwa, it is often omitted also in Swahili. A possible hypothesis is that the choice of OM vs. object drop in Swahili and lexical NP vs. object drop in Makhuwa might be linked to the importance of the referent within the rest of the paragraph. This would explain why in example (6b) the noun *ekofiyo* is repeated as a full NP, as it is relevant for the further discourse and will keep appearing throughout the continuation of the story. In contrast the object in example (5b) is a spoon which is not relevant for the continuing discourse as the episode ends right there.

Therefore, from this pilot study the preliminary analysis can be drawn that typological difference in object marking paradigms affects the discourse strategies of the two studied languages. The (non-)availability of OMs creates differences in the devices used for the tracking of referents in discourse. In Table 3 below the referent coding preferences for each of the languages are presented based on the texts analysed in the pilot study.

Table 3
Within-paragraph referents coding progressions

Swahili	lexical NP > OM > 0
Makhuwa (for non class nouns)	lexical NP > lexical NP > 0
Makhuwa (for class nouns)	lexical NP > OM

6. CONCLUSION

Interesting and unexpected patterns emerged from the activation status analysis in this pilot project, which call for the reconsideration of certain models generally accepted cross linguistically. In particular, the Makhuwa data revealed active given referents expressed by full NPs over and over again within only a couple of clauses. This goes against the prediction that the more accessible a topic is, the less phonological weight it tends to have (Givon 1983). The pilot study shows a possible link of this phenomenon to the importance of a referent in future discourse.

For an in-depth analysis of the link between object marking and discourse strategies in Bantu languages, a larger corpus of data in different genres needs to be studied. Moreover, additional parameters need to be included such as a more elaborated description of environments in which object marking and object drop can occur. A further development of this pilot study will continue to contribute towards a better understanding of the use of object marking in discourse as well as other referent tracking tools such as demonstrative pronouns. The latter were not included in the pilot study due to their low frequency in the analysed texts, but a deeper investigation of this topic would benefit from their inclusion. Lastly, a better understanding of text-structure in Bantu will further facilitate the future study of object marking in discourse.

REFERENCES

- Bearth, Thomas. 2003. Syntax. In: Nurse, Derek and Philippson, Gérard (eds.). *The Bantu Languages*. London: Routledge, 121-142
- Bresnan, Joan and Sam A. Mchombo. 1987. Topic, pronoun and agreement in Chichewa. *Languages* 63, no.4:741-782
- Bickel, Balthasar. 2003. Referential density in discourse and syntactic typology. *Language* 79, 708 – 736
- Chafe, Wallace, ed. 1980. *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, NJ: Ablex.
- Chafe, Wallace. 1994. *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago: University of Chicago Press.
- Duranti, A. 1979 . Object Clitic Pronouns in Bantu and the Topicality Hierarchy. *Studies in African Linguistics X*: 31-45.
- Givón, Talmy. 1983. Topic continuity in discourse: an introduction. In: Talmy Givon (ed.). *Topic continuity in discourse: a quantitative cross-language study* (Typological Studies in Language 3): 5-41. Amsterdam: John Benjamins.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2014. *Ethnologue: Languages of the World, Seventeenth edition*. Dallas, Texas: SIL International
- Marten, Lutz and Nancy C. Kula. 2012. Object marking and morphosyntactic variation in Bantu. *Southern African Linguistics and Applied Language Studies*, 30 (2): 237-253.
- Marten, Lutz. 2013. *Discourse-bound object drop in Yeyi as a Khoisan contact feature*. Presented at: SOAS Bantu workshop on word-order, information structure and language contact in Bantu, 1 Feb 2013.
- Nurse, Derek and Philippson, Gérard (eds.). 2003. *The Bantu Languages*. London: Routledge.
- Riedel, Kristina. 2009. *The Syntax of Object Marking in Sambia: A comparative Bantu perspective*. PhD dissertation. Leiden University.

- Seidl, Amanda H. and Dimitriadis, A. 1997. The discourse function of object marking in Swahili. *Papers from the regional meeting of the Chicago Linguistic Society* 33.
- Van der Wal, Jenneke. 2009. *Word order and information structure in Mkuwa-Enahara*. PhD dissertation. Leiden University.
- Van der Wal, Jenneke. 2010. Functions of demonstratives in Makuwa narratives. *Africana Linguistica* 16: 183-213.
- Van Staden, M., Senft, G., Enfield, N. J., & Bohnermeyer, J. 2001. Staged events. In: S. C. Levinson, & N. J. Enfield (Eds.). *Manual for the field season 2001*:115-125. Nijmegen: Max Planck Institute for Psycholinguistics.