EFFECTS OF HEAD-MARKING ON CONSTITUENT ORDER IN CHICHEWA*

Sam Mchombo
University of California, Berkeley.

INTRODUCTION

Standard analyses of the morphological structure of the verb in Chichewa, and other Bantu languages have the verb comprising a verb root (VR) to which verbal extensions such as causative, applicative, reciprocal, passive, stative, etc. are suffixed, and to which prefixes are added. The prefixes, analyzed as clitics in some studies, cf. Givón 1971, Mchombo (in press), include elements that encode information pertaining to agreement with the subject and the object, tense/aspect, negation, modality etc. The morphology of suffixation and prefixation (or cliticization) is sufficiently well established as to require no further comment. The following provides a typical example of such morphological organization:

1. Mkángo u-da-omb-án-its-á alenje ndí asodzi
   3-lion 3SM-pst-hit-recip-caus-fv 2-hunter and 2-fishermen
   ‘The lion made the hunters and the fishermen hit each other.’

In this example the VR –omb- ‘hit’ supports the reciprocal extension –an- and the causative extension –its-, the final vowel (fv) –a, as well as the subject marker u and the past tense marker da. This paper will focus on the subject and object markers, which appear on the verb head, and the consequence of their appearance on constituent order of the nominal phrases in the sentence.

* I am grateful to Andreas Kathol for suggesting the topic and to him, Adams Bodomo, Eyamba Bokamba, Joan Bresnan, Ron Kaplan, Al Mtenje, Steven Schaufele, and participants at the 6th International Lexical Functional Grammar Association conference held at the University of Hong Kong, June 2001, for comments on the material discussed in this paper. Preliminary versions of this material were delivered to audiences at the University of Sonora in Hermosillo, Mexico during a course on Bantu morphosyntax (May 2001) and at a colloquium held at El Colegio de Mexico in Mexico City, D.F. (May 2001). I am grateful to the participants, most especially to Isabel Barreras, Zarina Estrada Fernández, Maria Eugenia Vazquez Laslop, and Paulette Levy, and many others, for comments. I remain responsible for all errors.
HEAD-MARKING IN CHICHEWA

Chichewa shows both subject and object agreement in its verbal morphology. In finite verb forms the S[ubject] M[arker] is obligatory, while the single O[bject] M[arker] is optional. This is exemplified by sentence 2 below:

11 a. Njûchi zi-na-lúm-á alenje
   10-bees 10SM-pst-bit-fv 2-hunters
   ‘The bees hit the hunters’

   b. Zi-na-lúmá alenje njûchi.
   10SM-pst-bit-fv 2-hunters 10-bees
   ‘The bees hit the hunters’

In this sentence the SM zi in the verbal morphology contains the φ-features of the nominal njûchi ‘bees.’ Note that nominal njûchi whose φ-features are duplicated by the SM need not appear initially. It can appear in post-verbal position, for as long as it does not disrupt the string adjacency between the verb and its object. The object NP alenje ‘hunters’ must appear in the immediate post-verbal position, adjacent to the verb. The adjacency requirement is demonstrated by the ungrammaticality of sentence 3 below, in which the expression dzulo ‘yesterday’ is inserted between the verb and the object:

3 *Njûchi zi-na-lúm-á dzulo alenje
   10-bees 10SM-pst-bite-fv yesterday 2-hunters
This sentence would be well-formed if the expression dzulo ‘yesterday’ appeared after the object NP alenje ‘hunters.’

The verb can also be marked with an object marker (OM), an element that duplicates the φ-feature of the object NP. In the case of the nominal alenje, the relevant OM is wa. This is shown in sentence 4:

4  a. Njūchi zi-ná-wá-lum-a alenje.
   10-bees 10SM-pst-2OM-bite-fv 2-hunters

   ‘The bees bit the hunters.’

The presence of the SM and the OM has consequences on the word order requirements of Chichewa. For a start, the adjacency requirements between the verb and its object NP is relaxed and the ordering of the three major constituents of the sentences becomes free. Thus, the sentences in 4b-e below, showing the different ordering possibilities of the major constituents of the sentence 4a, are grammatical and have the meaning that the bees bit the hunters.

   b. SVO Njūchi zi-ná-wá-lum-a alenje
   c. OVS Alenje zi-ná-wá-luma njuchi
   d. VOS Zi-ná-wá-lum-a alenje njuchi
   e. SOV Njuchi alenje zi-ná-wá-lum-a

Further, the presence of the SM and OM license the omission of the actual noun phrases, as shown in (f) below:

   f. Zi-ná-wá-lum-a
   10SM-pst-2OM-bite-fv

   ‘They bit them’
In a detailed study of these morphological elements in Chichewa, Bresnan (1985), Bresnan & Mchombo (1986, 1987), and Mchombo (1984) analyzed these markers as incorporated pronominal arguments. The SM is ambiguously used for grammatical and anaphoric agreement whereas the OM is an incorporated pronominal argument. The analysis of the OM as an incorporated pronominal derives further support from the distribution in other Bantu languages such as Kikuyu (Bergvall 1986; Mugane 1997) or Lunda (Kawasha 2001). In those languages the OM is in complementary distribution with an overt nominal functioning as the object of the verb. Based on studies of Setawana, a dialect of Setswana, Demuth & Johnson (1989) analyzed the SM in that language as an incorporated pronominal which functions as the subject of the sentence. This is applicable to Chichewa on the pronominal argument analysis of the SM.

The seminal work of Bresnan & Mchombo focused on the status of the SM and OM, both of which are marked on the verbal head. Their comments on the effect of head-marking on constituent order was confined to the relative freedom of word order that accompanied the presence of the OM. This paper will take those observations a little further. Consider the following:

5 a. Njuchí izi zi-ná-lúm-á álenje
   10-bees 10prox.dem 10SM-pst-bite-fv 2-hunter
   awa ópúsa
   2prox.dem 2SM-foolish
   ‘These bees bit these foolish hunters.’

b. Mikángó i-tátu i-ná-gúmúl-á
   4-lions 4SM-three 4SM-pst-pull down-fv
   makólá ónse a-náyi
   6-corrls all 6SM-four
   ‘Three lions puled down all the four corrals.’
In these the subject and object NPs have internal constituents. Ordinarily, those constituents cannot be separated from the head noun. The sentences below are ungrammatical:

6  a. *Awa njuchí izi zi-ná-lúmá alenje
   2proxdem 10-bees 10proxdem 10SM-pst-bite 2-hunters
   ópúsa
   foolish

   b. *Awa ópúsa njuchí izi zi-ná-lúmá
      2proxdem foolish 10-bees 10proxdem 10SM-pst-bite
      alenje
      2-hunters

   Without the object marker the object NP cannot be discontinuous. On the other hand, in the presence of the OM not only does the order of the constituent words of the sentences become free but, it seems that the internal components of the NPs equally tolerate discontinuity.

7  a. Awa njuchí izi zi-ná-wá-lum-a
   2-prox-dem 10-bees 10prox dem 10SM-pst-2OM-bite-fv
   alenje ópúsa
   2-hunters 2SM-foolish

   b. Alenje zi-ná-wá-lum-a njuchí izi
      2-hunters 10SM-pst-2OM-bite-fv 10-bees 10-prox.dem
      awa ópúsa
      2-prox.dem 2SM-foolish
The data above show the possibility of discontinuity among the internal constituents of the NPs, induced by the presence of both SM and OM being marked on the verbal head. This kind of free word order as well as the possibility of syntactically discontinuous expressions has the hallmarks of non-configurationality. The question thus arises, is Chichewa a non-configurational language?

**THE CONFIGURATIONALITY PARAMETER**

Studies of Australian languages such as Warlpiri, initiated by Ken Hale, revealed properties about natural language that posed a problem for the conception of grammatical theory proposed within generative grammar. These languages display properties that were dubbed ‘non-configurational.’ The salient properties of non-configurational languages are:

i. free word order

ii. syntactically discontinuous expressions, and,

iii. null anaphora.

By syntactically discontinuous expressions is meant that non-adjacent nominals may correspond to a single verbal argument; and, null anaphora, according to Hale (1983) is the situation in which an argument (subject, object) is not expressed by an overt nominal in phrase structure.

Hale proposed to deal with the facts about Australian languages by separating lexical structure from phrase structure and positing a configurationality parameter for the application of the projection principle. In configurational languages, such as English, the projection principle was assumed to hold of the pair LS (lexical structure), PS (phrase
structure). In non-configurational languages, the projection principle holds of the LS alone.

In further analyses of configurationality, Jelinek posited the pronominal argument hypothesis (PAH) to account for the observed facts. The analysis reduces to the claim that the projection principle uniformly applies to lexical structure cross-linguistically. In configurational languages grammatical relations are marked by the order of the syntactic constituents, which satisfy the argument structure of the predicate. In non-configurational languages, null anaphors satisfy argument structure, hence, grammatical relations may be marked in the morphology as well as in the syntax. The nominal expressions that appear in the sentence are not verbal arguments but freely added expressions that may be referentially linked to pronominal arguments inside the verbal complex.

The pronominal argument hypothesis was reviewed in studies conducted within the framework of LFG (Austin & Bresnan 1996) and rejected as the optimal account for non-configurationality. Austin & Bresnan propose a dual structure hypothesis, where natural language is decomposable into linked parallel informational structures, such as constituent structure (c-s), functional structure (f-s), argument structure (a-s), etc, each of a different formal character. The grammar, within LFG “consists of a set of local, co-descriptive constraints on partial structures. There are no derivational or transformational operations involved: grammatical structures are defined by constraint satisfaction.” (Bresnan 1982).

Further, the parallel structures of LFG model different facets of the structure of language. The constituent structure, also called categorial structure, models the overt structure of forms of expression, encoding such surface structure relations as precedence and dominance, while the functional structure models the grammatical relations among syntactic functions. Argument structure deals with the grammatically expressible participants of eventualities (cf. Austin & Bresnan 1996). The various structures are associated by principles of functional correspondence, sometimes called linking or mapping principles.

The idea of a configurationality parameter is also rejected by Nordlinger (1998). Focusing on constructive case within Australian languages Nordlinger claims that “…there is no parametric distinction between configurationality and
nonconfigurationality. Rather, these two language types merely represent the extremes of a continuum: languages may identify grammatical relations in the syntax (fully configurational), in the morphology (fully nonconfigurational), or (more usually) by some mixture of the two.” (Nordlinger 1998:26). It seems that Chichewa lends some credence to this claim by displaying characteristics that place it somewhere between the extreme positions.

**CHICHEWA VERBAL MORPHOLOGY AND THE PRONOMINAL ARGUMENT HYPOTHESIS.**

The configurationality parameter of Hale rested on the separation of lexical structure from phrase structure and determination of whether the projection principle holds of both structures. By lexical structure Hale refers to predicates and their argument arrays which correspond to variables specified in the dictionary definition of a verb. Jelinek’s pronominal argument hypothesis derives from the claim that the argument variables are satisfied by pronominal arguments, which in the Australian languages may lack phonological realization. This account appears to be readily applicable to Bantu languages. Thus, consider sentence 4a again, repeated below as 8:

8 Njûchi zi-ná-wá-luma alenje

‘The bees bit the hunters’

It could be argued that the SM and OM are the pronominal arguments and the NPs njûchi ‘bees’ and alenje ‘hunters’ are not arguments of the verb. This is technically true. The fact that they can be omitted without inducing ungrammaticality testifies to that. Further, the complementarity between the OM and an overt nominal argument in some languages such as Gikuyu demonstrated the argument status of the OM. How then do the word order facts get and the discontinuous constituency of the nominal expressions get derived? Jelinek’s suggestion was to appeal to referential linking, treating the constituents
of the NPs as nominals that get linked to the incorporated arguments. This approach is rejected by Reinholtz, discussing comparable facts in Swampy Cree, on the grounds that “...the formation of discontinuous constituents does not depend on ‘referential linking’ and that it has the all the hallmarks of wh-movement in so called configurational languages” (Reinholtz 1999).

The determination of grammatical relations in Chichewa definitely appears to involve a mixture of syntactic order and morphological marking. The SM and OM as pronominal arguments marked on the verbal head are strictly ordered. The determination of grammatical relations based on their distribution involves ordering which may be linked to dominance relations. The OM is always sister to the verb stem, indicative of its status as an internal argument. It is in complementary distribution with an overt nominal, even in Chichewa where *inter alia* phonological cues such as tonal patterning, mark the overt NP as outside the VP configuration when the OM is present (see Bresnan & Mchombo 1987, Bresnan & Kanerva 1989). In brief, Chichewa has a VP configuration. It is the requirements of the VP that mandate the ordering relation between the verb and its internal argument. What is clearly the case is that the when the verb has a nominal phrase as its argument, the constituent integrity of the NP must be respected and retained, relaxed only when the OM is the argument. Why?

**DISCOURSE FUNCTIONS, GRAMMATICAL FUNCTIONS, AND CONSTITUENT ORDER**

Head-marking in Chichewa, just like pronominal argument marking, in Australian languages, has the result that the pronominal arguments bear the syntactic functions of subject and object. The nominal expressions that may appear are indeed extra-sentential and have discourse functions such as Topic. This is certainly the case with the NP that is linked to the OM. Such discourse elements have to be linked to elements bearing grammatical functions for the expression to be well-formed. This is where the theory of LFG provides the architecture for an elegant account of the linguistic facts. The theory posits constraints such as the completeness condition, coherence condition, extended coherence condition, which need to be satisfied. In the f-s grammaticized discourse
functions such as Topic (TOP) and Focus (FOC) will be included, with their linkage to the syntactic functions either through functional or anaphoric binding resolved. For instance the f-s of sentence 8 above may be represented as follows:

S -> NP, NP, VP
VP -> SM T/AOM-VS

Besides the anaphoric agreement that holds between the pronominal arguments and the NPs in the discourse structure, Bantu languages also display agreement between a nominal head and its modifiers. Consider the NPs alenje awa ópúsa ‘these foolish hunters’ and njuchí izi zópúsa ‘these foolish bees.’ In both, the form of the demonstrative ‘these’ is different, reflecting agreement with the head noun in number and gender (nominal class features). The same applies to the form of word ‘foolish’, which is derived from the verb –pusa ‘be foolish’ through the prefixation of the associative marker –a to which the class concord of the noun is applied. This yields the different forms of the word. Such agreement will be reflected in the f-s representation of the NP, where the constituents of the NP will have the same values for the specification of the gender and number attributes. The disparate bits of information associated with the constituents of the NPs are unified within the f-s which is where the anaphoric relation between the TOP nominal and the pronominal argument bearing the syntactic function is specified, satisfying the extended coherence condition. This is the condition that requires that all syntactic functions (including adjuncts and grammaticalized discourse functions) must be appropriately integrated into the f-structure (cf. Farsi-Fehri 1984; Nordlinger 1998). As such, the constituents of the NP do not have to be ‘referentially linked’ to the pronominal argument separately. The discontinuity of the constituents is possible in the c-s “[b]ecause of the many-one correspondence between the parallel c- and f-structures in violation of the projection principle, true discontinuous constituents are allowed. The principle of functional uniqueness, together with the free association of functions with constituents of the non-projective S node, can create a single functional constituent in f-structure corresponding to a ‘scattered’ set of c-structure nodes. Such a constituent will have the ‘merged’ interpretation. (Austin & Bresnan 1996:237).
THE LIMITS OF DISCONTINUITY

Although there is a degree of discontinuity among the constituents of the nominal expressions when the verbal head is marked with the subject and object pronominal arguments, there are limits on the possibilities of such discontinuity. Consider the following sentence:

9 Mkângo u-méné ú-ma-saká mbûzi
3-lion 3SM-relpro 3SM-hab-hunt10SM-goats
ú-ma-wa-sautsa alenje a-méné á-ma-gwetsá
3SM-hab-2OM-bother 2-hunters 2SM-relpro 2SM-hab-fell
mitêngo
4-trees
‘The lion which hunts goats bothers the hunters who fell trees’

The formation of relative clauses in Chichewa is, largely, comparable to that of English. Chichewa is a head-initial language and, within relative clauses, the relative clause follows the head noun. There is a relative marker –mene which is marked for agreement with the head noun and introduces the relative clause. If the relativized Nominal is the object of the verb, the OM is optionally absent, but normally present, functioning as a resumptive pronoun. The presence of the relative marker –mene also has the phonological effect of marking the verb within the relative construction with a high tone. Consider the following:

10 a. mkângo u-ku-sáká mbûzi
3-lion 3SM-pres-hunt 10-goats
‘The lion is hunting goats’
In these examples the tone patterns on the verb *ukusaka* ‘it is hunting’ are different, in part because of the presence of the relative marker in (b). The fact that the tone marking functions as a phonological cue of the different constructions makes it possible for the relative marker *u-mene* to be deleted or dropped. Thus construction 10c below, which only differs tonally from sentence 10a above, is still construed as a relativized NP configuration:

c. mkángó  ú-kú-sáká mbûzi
   3-lion  3SM-pres-hunt 10-goats
   ‘The lion which is hunting goats.’

Returning to sentence 9 above, the verb *sautsa* ‘bother, trouble’ is marked with the SM *u* agreeing with *mkángo* ‘lion’ and the OM *wa*, agreeing with *alenje* ‘hunters.’. While the order of the nominal expressions *mkango u-mene u-ma-saka mbuzi* ‘the lion which hunts goats’ and *alenje a-mene a-ma-gwetsa mitengo* ‘the hunters who fell the trees’ is free, the constituents of those nominal expressions cannot be discontinuous. The following sentence is, at best, questionable. :

11 ?*alenje* mkángó ú-ma-*wa*-sautsa a-méné
   2-hunters 3-lion 3SM-hab-2OM bother 2SM-relpro
   á-ma-gwetsá miténgo u-méné ú-ma-saká mbûzi
   2SM-hab-fell 4-trees 3SM-rekpro 3SM-hab-hunt 10-goats
   ‘The hunters the lion bother them, who fell trees, that hunts goats.’

The nominal expressions are both relativized NPs. The possibility of extraposing the relative clause does not completely rescue the sentence from ungrammaticality. The object NPs within those relative clauses cannot be moved out largely because the verbal head is itself not marked with the OM. Now consider the following where the verbal heads within the relative clauses are marked with the OMs:
Mkángo uméné úmazisaka mbûzi úmawasautsa alenje améné amaigwetsa miténgo
‘The lion which hunts the goats bothers the hunters that fell the trees’

This sentence does not allow for the range of possible word orders that are normally associated with head-marking in Chichewa. Thus,

a. *Mkángo mbûzi alenje miténgo úmawasautsa umêné umazisaka améné ámaigwetsa
The lion, the goats, the hunters, the trees, it bothers them which hunts them (goats) who fell them (trees)

b. *Mbûzi mkángo uméné úmazisaka miténgo alenje améné amaigwetsa úmawasautsa
The goats the lion which hunts them the trees the hunters who fell them (the trees) it (the lion) bothers them (hunters)

Part of the problem associated with 11b is that it seems to have mbuzi ‘goats’ as its topic yet the sentence has to do with the fact that the lion which hunts the goats bothers the hunters who fell trees. In brief, when the nominal expressions have relative clauses, discontinuity among the constituents is more difficult, certainly reducible to island effects.

In her study of discontinuous constituents in Swampy Cree Reinholtz (1999) proposes to treat discontinuous constituents as the output of a Focus mechanism, which “…picks out a nominal modifier belonging to a larger NP, and places this in a preverbal Focus position where it is separated from the noun it qualifies by other material” (Reinholtz 1999: 28). She proceeds to adopt the analysis of Swampy Cree proposed in Russell & Reinholtz (1996) that discontinuous NPs involve movement, which moves the movement of a nominal modifier out of a containing NP to a pre-verbal focus position. This movement-based approach leads Reinholtz to assimilate such discontinuous constituents to being an instance of Wh-movement. She notes that movement to Focus position is
commonly grouped with wh-movement and the discontinuous constituents in Swampy Cree share common characteristics with wh-movement. Thus, they show the ability to span several clauses, cannot move material out of adverbial constituents, and have limited application in relative clauses or embedded questions.

It is an open question as to whether discontinuous constituents can be assimilated to wh-movement. Certainly failure to affect relative clauses can be handled in terms of island effects. Even if an analysis along the lines of wh-movement were to be adopted, within LFG there are no movement rules. The resolution of wh-effects can be handled in terms of functional uncertainty. The island effects associated with relative clauses are equally amenable to analysis in terms of functional uncertainty. Significantly, it is not the case that movement needs to be invoked since the domain where functional uncertainty applies is the f-s. The c-s to f-s mapping algorithm allows for the possible scattering in c-s of information which, in the f-s, is unified. Limits on discontinuity are the result of satisfying constraints on partial structures of parallel informational structures.

**CONCLUSION**

Head-marking in Chichewa results in discontinuous constituency of the nominal expressions construed as dependent on the incorporated pronominal arguments. While this may appear to accord languages like Chichewa the appearance of being non-configurational, the language seems to have a VP. The incorporated pronominal arguments also have a fairly rigid order, characteristic of clitics in Bantu languages in general. In this regard, the clitics differ from the verbal extensions or suffixes, which affect argument structure and are subject to an array of linguistic processes excluded from the pre-VS clitic domain (cf. Mchombo, in press). The verbal suffixes are subject to variable order, within limits. The rigidity of the clitics captures c-command relation (borrowing the terminology of the principles and parameters theory) between the incorporated pronominal arguments. Thus Chichewa offers an example of a language that is configurational yet allows for discontinuous constituents under the specific conditions of head-marking, comparable to non-configurational languages. This suggests the independence of discontinuity of constituents from non-configurationality, as noted by
Austin & Bresnan, as well as Nordlinger. It still remains to provide an exact specification of this and of the limits on discontinuity in Chichewa within the theory of LFG.

REFERENCES.


