Linear Order Constraints on Split NPs in Chichewa[*]

Sam Mchombo
University of California, Berkeley

This paper focuses on restrictions on the ordering of internal constituents of noun phrases in Chichewa, especially when those constituents are discontinuous. The motivation for discontinuity of the NP constituents will be given, together with discussion of constructions that can be subsumed under this rubric but that do not really involve discontinuity in the canonical sense. These are constructions where a topic NP in a left periphery position is either linked anaphorically with a modifier "remnant" or semantically with its hyponym in post-verbal position. According to Guthrie's classification of Bantu languages, Chichewa is placed in zone N unit N31. It is regarded as a dialect of Nyanja, classified as belonging to unit N30 (Guthrie 1967-71).

1   Introduction

Recent work on the interaction of discourse structure and morphosyntax in Chichewa has brought to light the fact that the language is partially non-configurational (Mchombo 2002, 2003).[1] Chichewa has the verbal morphology that is standard for Bantu languages where the verb comprises, *inter alia*, the

---

[1] According to Greenberg (1963) the languages of Africa are classified into four families: Niger-Kordofanian, Nilo-Saharan, Afroasiatic and, Khoisan. Niger-Kordofanian is the family to which the vast majority of the languages of sub-Saharan Africa belong. One large subdivision of the Niger-Kordofanian language family is the Niger-Congo. The major language group in this subgroup, to which the majority of the languages spoken in the southern half of the African land mass belong, is the Bantu language group. Bantu languages are spoken by nearly a third of Africa's total population.

subject marker (SM) and object marker (OM) that cross-reference nominals that are essentially internal topic (I-TOP) and external topic (E-TOP) respectively (Baker 2003; Morimoto 2000). This is illustrated by sentence 1 below:

(1)   Mkángo   u-da-wá-ómb-án-íts-a   alenje   ndí   asodzi
        3lion   3SM-pst-2OM-hit-recip-caus-fv 2hunters   and   2fishermen
        'The lion made them hit each other, the hunters and the fishermen'[2]

To the verb root **-omb-** 'hit' is attached the SM **u-** and OM **-wa-** that agree with the NPs **mkángo** 'lion' and the coordinate **NP alenje ndí asodzi** 'hunters and fishermen' respectively.

The subject marker (SM) and object marker (OM) have received sustained attention in recent studies. In Chichewa, the OM has been analyzed as an incorporated pronominal argument whereas the SM has been claimed to be functionally ambiguous between agreement marker and incorporated pronominal (cf. Bresnan & Mchombo 1987; Mchombo 2004). The presence of especially the OM has been correlated with free word order of the constituents of the sentence. Further, it is correlated with discontinuity of the constituents of the TOPIC nominals when those are internally complex (Mchombo & Morimoto 2004). This paper will focus on constraints on the linear order of such discontinuous constituents in Chichewa.

2   Discontinuity of constituents of the NPs

A significant aspect of the presence of the SM and, especially, of the OM in the verbal morphology is the discontinuity of the constituents of an internally complex dislocated Topic NP. This points at the possibility of Chichewa being partially non-configurational. Consider the examples below:

---

[2]   The following abbreviations are used here:

appl	applicative
assoc	associative marker
caus	causative
fv	final vowel
hab	habitual
perf	perfective marker
pres	present tense
proxdem	proximal demonstrative
pst	past tense
recip	reciprocal
relpro	relative pronoun

(2) a. *Njuchí   izi   zi*-ná-**wá**-lúm-á   **álenje**   awa
       10bees   10prox.dem 10SM-pst-2OM-bite-fv   2hunter   2prox.dem
       ópúsa
       2SMfoolish
       'These bees bit these foolish hunters.'
   b. **Awa**   *njuchí   izi*   *zi*-ná-**wá**-lum-a   **alenje**
       2prox.dem   10bees   10prox.dem   10SM-pst-2OM-bite-fv   2hunters
       ópúsa
       2SMfoolish

In this example, the proximal demonstrative **awa** 'these', agreeing in ɸ-features with the head noun **alenje** 'hunters', appears discontinuous with the rest of the phrase. This is disallowed when the OM is absent from the verbal morphology. Thus sentence (2c) below, in which the proximal demonstrative is fronted, but the OM is not present in the verbal morphology, is ungrammatical.

   c. *__Awa__   *njuchí   izi*   *zi*-ná-lúm-á   **alenje**
       2prox.dem   10bees   10prox.dem   10SM-pst-2OM-bite-fv   2hunters
       ópúsa
       2SMfoolish

The analysis of the SM and OM as pronominal arguments fits Chichewa into the pronominal argument hypothesis (Jelinek 1984). Basically it is a common trait of polysynthetic languages to have missing nominals, allowing morphological elements to provide the relevant information. A central concern in analyses of polysynthetic languages has been the search for explanation of the connection between such highly articulated systems of argument agreement marking, and the free ordering and omission of independent nominal expressions. According to Phillips (1993) an answer to this question was offered by Wilhelm von Humboldt in 1836. In a study of Nahuatl, Wilhelm von Humboldt claimed that "…it is the affixes on the verb which fill the argument positions of the verb, and that independent expressions are only loosely linked to these affixes" (Phillips 1993: 173).

Within generative grammar the issue has been revived by Jelinek's analysis of Warlpiri. Jelinek focused on why it is that independent expressions cannot appear in argument positions in polysynthetic languages. Her suggestion, couched in her pronominal argument hypothesis (PAH), is that the affixes are assigned thematic roles and Case. As such, these are unavailable to license independent phrases in argument positions. Jelinek's analysis of pronominal affixes as arguments got extended to account for the property of non-

configurationality in polysynthetic languages. Hale (1983) characterized such languages as manifesting the following properties:

(i)   free word order
(ii)   syntactically discontinuous expressions, and
(iii)   null anaphora.

By "null anaphora" Hale referred to "the situation in which an argument (e.g., subject, object) is not expressed by an overt nominal expression in phrase structure" (Hale 1983: 40). By syntactically discontinuous expressions is meant the situation where non-adjacent constituents may correspond to a single argument of the verb, resulting in discontinuous expressions.

The facts indicated above appear to warrant the classification of Chichewa as (partially) non-configurational. However, although free word order and possible omission of the nominals are intimately connected with the presence of the SM and OM, as the example above illustrate, discontinuity of the phrasal constituents is also possible in the absence of the OM. In other words, the incorporated pronominal arguments, especially the OM, constitute necessary but not sufficient conditions for discontinuity of constituents. Some of the data will be examined here.

2   Linearization and the relevance of anaphoric linkage

The case to be reviewed here does not fit the technical characterization of split NP. The situation pertains to structures where the NPs that are correlated are in a relation reducible to anaphora. In such cases there is a relation between a topic NP in the left periphery of the clause and a 'remnant' i.e., the element in the post-verbal position. The relation is basically anaphoric in that the remnant is in an appositional relation to the dislocated or fronted head noun.[3] In such cases there are constraints on word order too. These require that the head noun be in left-dislocated position. Consider the following:

---

[3]   The use of the term "fronted", along with other terminology traditionally employed in derivational theories of grammar is merely for convenience, used here simply as labels. The use does not either amount to subscription to any derivational theory of grammar.

(3) a. *Makású* <u>mbidzí</u> <u>zi</u>-na-**í**-gúl-íl-á   *awa*   *óbúntha*
        6hoes   10zebras   10SM-pst-4OM-buy-appl-fv   6these   6SMblunt
        mikángó   **yókálamba**
        4lions   4SMaged
     '(As for) hoes, the zebras bought (for) the aged lions these blunt ones.'

In this example, the OM **-i-** agrees with **mikángó yókálamba** 'aged lions', not with **makású awa óbúntha** 'these blunt hoes'. The split NP has its head noun **makásu** 'hoes' appearing sentence-initially. The modifier remains in post-verbal position maintaining an anaphoric or appositional relation between it and the head noun. That the remnant must be construed as anaphoric to the fronted noun is further exemplified by the following:

   b. *Mbûzi*   <u>atsíkáná</u>   á   mfúmu   <u>a</u>-a-gul-á   *zákúd-a*
       10goats   2girls   2AssocM   9chief   2S-perf-buy-fv   10SM-black
     Lit. 'Goats, the chief's girls have bought black (ones).'

   c. *Mikánda*,   <u>anyaní</u>   ópúsa   <u>a</u>-na-b-á   *yófíila*
       4beads   2baboons   2SMsilly   2SM-pst-steal-fv   4SMred
     Lit. 'Beads, the silly baboons stole red ones.'

In these examples, the modifiers **zákúda** 'black' and **yófíila** 'red' agree in the relevant gender and number features with the head nouns **mbûzi** 'goats', and **mikánda** 'beads', respectively. They are also construed as involving a null pronominal that is anaphorically bound by the head noun. The constraints on word order, requiring the modifier to be the remnant and ordered after the head noun rather than preceding it can be shown by the ungrammaticality of the following:

(4) a. **zakuda*   <u>atsíkáná</u>   á   mfúmu   <u>a</u>-a-gul-á   *mbûzi*
        10SMblack   2girls   2AssocM   9chief   2SM-perf-buy-fv   10goats

   b. **yofiila*   <u>anyaní</u>   ópúsa   <u>a</u>-na-b-á   *míkánda*
       4SMred   2baboons   2SM-silly   2SM-pst-steal-fv   4beads

Chichewa is a head initial language and, within a Noun Phrase (or Determiner Phrase (DP)), the modifiers are ordered after the head noun. The constraint on word order when discontinuity is not induced by the presence of the OM seems to require preservation of the 'base' order. Such constraints on linearization of discontinuous constituents has been observed in other languages. For instance,

in the Algonquian language of Ojibwe (cf. Kathol & Rhodes 1999) it has been noted that linear order of discontinuous elements preserves the 'base' order.

The preservation of the base order in the output (split) structure undermines a movement analysis of split NPs. The movement would need to be sensitive to the order of the moved constituents, giving the movement rule global power.

The constraints on linearization of split NPs in Chichewa seems reducible to command relations in anaphora. Thus, the remnant must be c-commanded by the antecedent, preserving in the constituent structure (c-s) representation, functional-command relations, indicated in the functional structure (f-s), adopting the architecture and terminology of the theory of lexical-functional grammar (LFG) (cf. Bresnan 2001; Dalrymple 2001).[4] The theory of LFG does not give recognition to movement operations to account for discontinuity of elements or construal of a 'displaced' element with the 'base' position.

Although there will not be a technical exposition of the theory of LFG here, in broad outline, the theory maintains that languages have an overt organization, modeled by the constituent structure (c-s), that is variable, within limits, across languages. Languages also have an internal organization, modelled by the functional structure (f-s) that is largely invariant. The f-s is shown in the form of an attribute-value matrix (AVM). This represents functional information such as subject, object, as well as information pertaining to the resolution of agreement, conditions on binding relations, functional identification of discourse elements with phonologically null governed grammatical functions, etc. A subcomponent of that is the argument structure (a-s) that deals with lexical argument structure and the relation between arguments and grammatical functions. Another component is that of discourse structure (d-s) in which such information structure notions as TOPIC and FOCUS feature, and their relation with grammatical functions are stated. The relation between the various structural representations is non-derivational. Rather, there is constraint satisfaction such that pieces of c-s are associated with pieces of f-s in systematic ways. Consider the representation of the English sentence "The baboon read the book". The c-s structure of this sentence is represented by the phrase structure configuration given below, whereas the f-s is represented by the AVM indicated.

---

[4]   The theory of lexical-functional grammar (LFG) is presupposed or adopted in this work. The technical aspects of the theory are not discussed for reasons of space. Further, theoretical expositions of the theory are, currently, readily available.
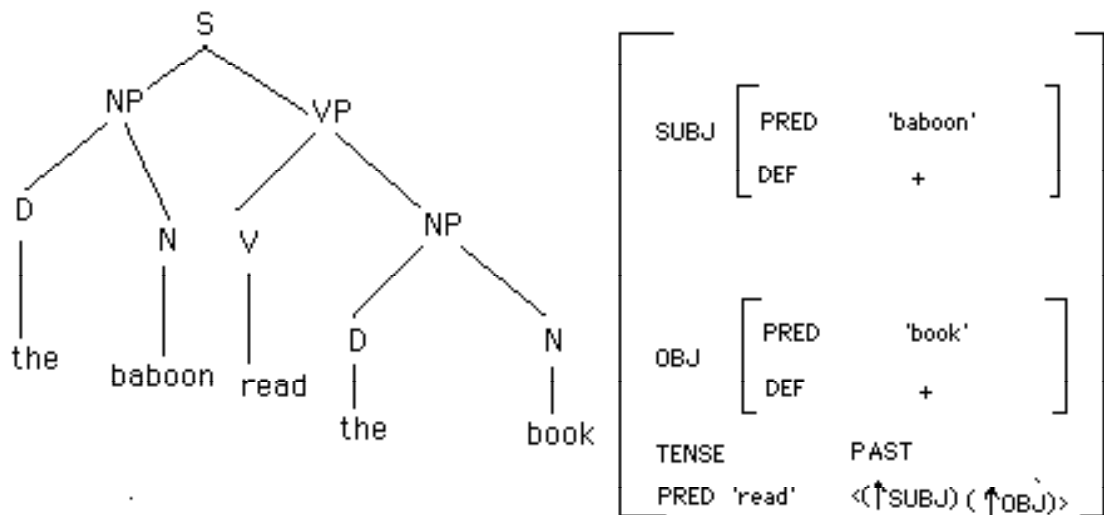
Figure 1

In this English sentence the SUBJ function is associated with the NP directly dominated by the S node in the c-s. The OBJ is associated with the NP that is the daughter of the VP. The tense feature has zero morphology on the verb. While abstracting away from the technical details of the mapping algorithm of the f-s to the c-s it should be noted that the mapping is, as indicated, non-derivational.

3   Super-ordinate/hyponomy relations

The situation of apposition or anaphoric relation indicated above is paralleled by that where the dislocated NP specifies a super-ordinate term with which the remnant is in a subset or hyponym relation. Technically, this cannot be subsumed under discontinuity in the strict sense in that the related components arguably do not form a constituent at any level of linguistic representation. Consider the examples below:

(5) a. *Nsomba*,   <u>anyání</u>   ópúsa   <u>a</u>-na-gúl-á   *milämba*
        10fish   2baboon 2SMfoolish 2SM-pst-buy-fv 4mudfish
      '(As for) fish, the foolish baboons bought mudfish.'
     b. *Miténgo*,   <u>njovu</u>   z-anu   <u>zi</u>-ku-gwéts-á   *milámbe*
         4trees   10elephants 10SM-your 10SM-pres-fell-fv 4baobabs
      '(As for) trees, your elephants are pulling down (the) baobabs.'

As in the previous cases where the head noun preceded the remnant in apposition or in an anaphoric relation with it, the super-ordinate term in this

latter case is required to precede the hyponym. The super-ordinate term appears in a left-dislocated position while the hyponym appears in the post-verbal position. The sentences are ungrammatical if the hyponym precedes the super-ordinate term even when the construction involved is, arguably, a right dislocation, as below:

(6) a. *Milämba*   <u>anyaní</u>   ópúsa   <u>a</u>-na-gúl-á   *nsómba*
          4mudfish   2baboon   2SMfoolish   2SM-pst-buy-fv   10fish
     b. *Anyaní*   ópúsa   a-na-gúl-á   *milämba*   *nsómba*
          2baboons   2SMfoolish   2SM-pst-buy-fv   4mudfish   10fish

This situation is atypical of genuine discontinuous constituents in the sense that the dislocated topic and the remnant do not form a constituent at any level of linguistic representation (cf. Nolda 2000) for relevant observations about German). The necessary condition is that the super-ordinate/hyponym relation must hold between the relevant noun phrases. Secondly, the presence of the OM is disallowed. Thus the sentences below are ungrammatical:

(7) a. *Nsómba,*   <u>anyání</u>   ópúsa   <u>a</u>-na-*zi*-gúl-á   *milämba*
          10fish   2baboon   2SMfoolish   2SM-pst-10OM-buy-fv   4mudfish
        '(As for) fish, the foolish baboons bought (them, i.e. fish) mudfish.'
     b. *Nsómba,*   anyání   ópúsa   a-na-*i*-gúl-á   *milämba*
          10fish   2baboon   2SMfoolish   2SM-pst-4OM-buy-fv   4mudfish
        '(As for) fish, the foolish baboons bought (them) mudfish'

In (7a) the OM agrees with the (external) Topic NP **nsómba** 'fish'. The presence of the OM is a necessary condition for the NP to assume the role of external topic (Morimoto 2000) and is consistent with the pronominal argument status of the OM (cf. Bresnan & Mchombo 1987; De Cat 2002). The ungrammaticality of the sentence seems to derive from "overcrowding' of nominal expressions. The OM is a pronominal argument anaphorically linked to the external topic but then the remnant is another argument of the verb, comparably anaphorically linked to the E-TOP.

In (7b) the OM agrees with the remnant or hyponym expression **milämba** 'mudfish'. This is also ungrammatical but, ostensibly, for different reasons. The OM makes the remnant a TOPIC and, somehow, the TOPIC to TOPIC relation between the super-ordinate term and the hyponym that ensues makes for ungrammaticality or unacceptability.

The constraints applying to the order of the super-ordinate Topic NP in a left-dislocated position and the hyponym remnant extends to quantified NPs (QNP).

4   Quantified NPs and Discontinuity

When a QNP is discontinuous, the Topic head noun must be in a dislocated position and must precede the quantifier. This is illustrated by the following:

(8) a. *Zisakasa*   <u>mkángó</u> <u>wópúsa</u> <u>u-na-máng-á</u>   *zi-sanu*   *z-ókha*
          8grass.hut 3lion   3SMsilly 3SM-pst-build-fv 8SM-five 8SM-only
        Lit. 'Grass huts, the silly lion built only five.'
      b. *Zisakasa*   <u>mkángó</u> <u>wópúsa</u> <u>u-na-máng-á</u>   *zó-chépa*
          8grass.hut 3lion   3SMsilly 3SM-pst-build-fv 8SM-few
        Lit. 'Grass huts, the silly lion built a few.'

The precedence relation between the head noun and the remnant or the quantifier is underscored by the following sentences where the quantifier precedes the head noun:

(9) a. **Zi-sanu   z-ókha*   <u>mkángó</u> <u>wópúsa</u> <u>u-na-máng-á</u>   *zisakasa*
          8SM-five 8SM-only 3lion   3SMsilly 3SM-pst-build-fv 8grass.hut

     b. **Zó-chépa* <u>mkángó</u> <u>wópúsa</u> <u>u-na-máng-á</u>   *zisakasa*
          8SM-few 3lion   3SMsilly 3SM-pst-build-fv 8grass.hut

As in the case of the ordering relations in the super-ordinate/hyponym relation, the presence of the OM does not make for marked improvement in grammaticality. In general, the presence of the OM with quantified expressions is subject to various restrictions. It might be expected that the OM, as an incorporated pronominal argument, should be bound by a referring expression. QNPs are non-referring, a characteristic that constituted motivation for the rule of Quantifier Raising (QR) to derive logical form within the principles and parameters theory (PPT). The argument reduced to the claim that as non-referring expressions QNPs could not be assigned θ-role. To avoid violation of the θ-criterion, the quantifier was raised to a non-argument (A-bar) position, leaving a variable in the argument position (May 1985; Neale 1996). The argument here is that the OM should not cross-reference a QNP (cf. Deen 2004; De Cat 2002).[5] However, since quantifiers bind variables, and pronouns can

---

[5]   Discussion of the relation between quantified noun phrases and the subject marker and object marker in Bantu languages is beginning to gain attention. Kamil Ud Deen has raised important questions in some of his work. The questions are likely to be focal points of future research in Bantu linguistics.

function as bound variables (cf. Higginbotham 1980) bound variable interpretation is allowed between the OM and a quantified NP, appropriately in dislocated position. Consider the following:

(10) *Mbidzí   zi-sanu*   mkángó u-ku-*zí*-thámángits-á   mwá
     10zebra   10SM-five   3lion   3SM-pres-10OM-chase-fv   18AssocM
     nkhanza
     10-cruelty
     Lit. 'Five zebras, the lion is chasing them in a cruel manner (cruelly).'

The relation between QNPs and incorporated pronominal arguments is clearer in the case of the SM. The subject marker is obligatory and is marked on all the modifiers of the head noun as well as on the verb. Such obligatoriness calls for variable treatment of the SM and OM in Chichewa and other Bantu languages (cf. Deen 2004; Letsholo 2002).

This underlies the variable treatment of the NPs that are cross-referenced by these markers. Thus, while the NP cross-referenced by the OM clearly has a TOPIC function, is extra-sentential and, in some Bantu languages it cannot appear within the same simple clause as the OM. This is attested in Kikuyu, a Bantu language spoken in Kenya (cf. Mugane 1997), the NP that is cross-referenced by the SM is, in many cases, required and is structurally close to the SM. In Kinande, another Bantu language, when the OM is present in the verbal morphology, the NP that is cross-referenced by the OM must be in a left-dislocated position and ordered to the left of the NP that is cross-referenced by the SM (Baker 2003). Adopting a version of the pronominal argument hypothesis for Bantu languages, and factoring in these noted asymmetries, Morimoto (2000) has analyzed the NPs as discourse licensed Topic elements (I-TOP and E-TOP). The interaction between such information structure elements and morphosyntactic organization is an issue of current interest.

Returning to the issue of constraints on linear order noted above, the observations seem to underscore the relevance of linear precedence relations to semantic interpretation and the need for articulation of linearization principles. Alternatively, this can be viewed as a constraint on the preservation or transparency in c-structure of f-structure relations. For instance, while command relations are stated in f-s, the mapping of the elements in c-s should make those relations transparent, facilitating the computation of the construal of the expression. Take, by way of example, the super-ordinate/hyponym relation and the overt realization of the elements related in that fashion. The relation is that of set inclusion. The super-ordinate term properly includes the hyponym, such that negation of the super-ordinate term entails negation of the hyponym. In brief, if something is not a fish, then it could not be a mudfish or salmon. The set

relation of inclusion may then be overtly indicated by the linearization that makes the super-ordinate precede the hyponym.

5 Limits on Discontinuity

In the previous sections attention has been on elements that are discontinuous, in some specifiable sense. Here, attention will turn to cases where discontinuity of the constituents of the NP is not possible even when the OM is present.

5.1 Relative clauses

One case where discontinuity of the constituents is not possible involves relative clauses (cf. Mchombo 2003, 2004). This is shown below:

(11) Mkángo   u-méné   ú-ma-sáká   mbûzi   ú-ma-*wa*-saútsa
       3lion   3SM-relpro 3SM-hab-hunt 10goats 3SM-hab-2OM-bother
       alenje   *a-méné*   *á-ma-gwetsá*   *mitêngo*
       2hunters 2SM-relpro 2SM-hab-fell   4trees
       'The lion which hunts goats bothers the hunters who fell trees.'

In Chichewa relative clause formation there is a relative marker **-méne**, which introduces the relative clause. It is marked for agreement with the relativized noun. When the relativized nominal is the object of the verb, the OM is optionally present, functioning as a resumptive pronoun.

    Phonologically, the verb within the relative clause introduced by **-méne** is marked by a high tone. Comparable observations have been made about relative clauses in Lunda, a language spoken in West Central Africa, particularly in the north east part of Angola, south-western region of the Democratic Republic of the Congo, and north-western part of Zambia (Kawasha 1999a, 1999b). Returning to Chichewa, consider the following:

(12) a. mkángo   u-ku-sáká   mbúzi
          3lion   3SM-pres-hunt   10goats
    'The lion is hunting goats.'
     b. mkángó   u-méné   ú-kú-sáká   mbúzi
          3lion   3SM-relpro   3SM-pres-hunt   10goats
    'The lion which is hunting goats...'

In these examples the tone patterns on the verb **ukusaka** 'it is hunting' are different depending on the presence or absence of the relative marker. The exploitation of tone to signal a relative construction licenses the omission the

relative marker **u-méné**. Sentence (12c) below, which only differs tonally from sentence (12a) above, is construed as a relativized NP configuration:

    c.   mkángó   ú-kú-sáká   mbúzi
         3lion   3SM-pres-hunt 10goats
   'The lion that is hunting goats...'

Returning to sentence 11 above, the verb **sautsa** 'bother, trouble' is marked with the SM **u**, for class 3, agreeing with **mkángo** 'lion'. The OM **wa** agrees with **alenje** 'hunters'. While the order of the nominal expressions **mkángó u-méné ú-ma-sáká mbûzi** 'the lion which hunts goats' and **alenje a-méné á-ma-gwétsá miténgo** 'the hunters who fell the trees' is free with respect to the verb, the constituents of the nominal expressions cannot be discontinuous. The following sentence is, at best, questionable:

(13) ?**Alenje**   *mkángó*   *ú*-ma-**wa**-saútsa   **a-méné**   **á-ma-gwetsá**
       2hunters   3lion   3SM-hab-2OM bother   2SM-relpro 2SM-hab-fell
       miténgo   *u*-méné   *ú*-ma-saká   mbúzi
       4trees   3SM-relpro 3SM-hab-hunt   10goats
       'The hunters the lion bother them, who fell trees, that hunts goats.'

The nominal expressions are relativized NPs. The possibility of extraposing the relative clause, possible in Chichewa, does not completely rescue the sentence from ungrammaticality. The object NPs within those relative clauses cannot be moved out in part because the OM is excluded from the verbal morphology. Now consider sentence 14 below. In this one, the relativized NP configurations, functioning as subject as well as the one in the post-verbal position both have their verbal heads within the relative clauses marked with resumptive OMs:

(14) Mkángo uméné úma**zi**saka **mbûzi** ú-ma-wa-saútsa alenje améné
      ama*i*gwetsa *miténgo.*
      'The lion which hunts the goats bothers the hunters that fell the trees.'

In addition, the main verb also has the OM in its morphological composition. Yet this sentence does not allow for the range of possible word orders that are normally associated with head-marking in Chichewa. Neither one of the sentences below is grammatical:

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

The problem here could be attributed to the multiplicity of nominal expressions with similar discourse functions, making for processing or information structuring difficulties. In general, when the nominal expressions have relative clauses, discontinuity among the constituents is more difficult. The problem could equally be reducible to island effects suggesting the presence of *wh*-movement, at any rate at the level of LF within PPT. Prior analyses have subsumed relative clauses under *wh*-movement (Chomsky 1977; Ngonyani 1999). In Chichewa relative clause formation need not be dealt with in movement terms (cf. Mchombo & Mtenje 1983). The presence of the OM functioning as a resumptive pronoun, is not correlated with movement.

5.1.1 Recursive noun phrases

Another case that restricts discontinuity involves recursive constructions with associative phrases. This is exemplified by NPs with a head noun and a complement introduced by the associative marker **-a** 'of' as shown below:

(16) a. chipanda　　chá　　　　kazitápe
　　　　　　7calabash　7AssocM　1a.spy
　　　　'the spy's calabash'
　　　　Lit. 'calabash of spy'

This construction allows for recursion, shown in the following construction:
　　　b. chipanda　　ch-á　　　　kazitápé　w-á　　　　　alenje
　　　　　　7calabash　7-AssocM　1a.spy　　1SM-AssocM　2-hunters
　　　　'calabash of the spy of the hunters'

In the sentence below the φ-features of this noun phrase are duplicated by the OM in the verbal morphology.

(17) anyaní   á   mísala   a-ku-**chí**-phwány-a
     2baboons 2AssocM   4madness   2SM-pres-7OM-smash-fv
     chipanda chá   kazitápé   wá   alenje
     7calabash 7AssocM   1a.spy   1AssocM   2hunters
     'The mad baboons are smashing the calabash of the hunters' spy.'

While re-ordering of the nominal expressions that are the major constituents of the sentence does not pose any problems, discontinuity of the constituents of the nominal expressions is tolerable only within limits. Consider the following:

(18) a.   Chá   kazitápé   wá   alenje
          7AssocM   1a.spy   1AssocM   2hunters
          anyaní   á   mísala   a-ku-chí-phwány-a   chipanda
          2baboons 2AssocM 4madness   2SM-pres-7OM-smash-fv   7calabash
          Lit. 'Of the spy of the hunters the mad baboons are smashing the calabash.'

This sentence is grammatical, with the complement of the head noun **chipanda** 'calabash' fronted. The results are radically different when the fronted constituent is the more deeply embedded complement of **kazitápe** 'spy', viz. **wá alenje** 'of the hunters.' In such cases the result is ungrammatical.

     b. ****Wá   alenje**   anyaní   á   misala
          1AssocM 2hunters   2baboons 2AssocM   4madness
          a-ku-**chí**-phwány-a   **chipanda   chá   kazitápe**
          2SM-pres-7OM-smash-fv   7calabash   7AssocM   1a.spy

The clash in the φ-features of the displaced constituent and of the OM does not account for the ungrammaticality since when that is controlled for, the results remain bad. Consider the sentence below where the φ-features of the OM do not clash with those of the fronted complement associative phrase.

     c. Anyaní   á   mísala   a-ku-**chí**-phwány-a
          2baboons 2AssocM 4madness   2SM-pres-7OM-smash-fv
          chipanda   chá   chiphadzúwá   chá   alenje
          7calabash 7AssocM   7beauty.queen 7SMAssocM   2hunters
          'The mad baboons are smashing the calabash of the beauty-queen of the hunters.'

Moving the Associative Phrase 'of the hunters' as in example (18b) still yields ungrammatical results despite agreement between the OM and the moved Associative Phrase.

d. ***Chá**   **alenje**   anyaní   á   misala
    1AssocM   2hunters   2baboons   2AssocM   4madness
    a-ku-**chí**-phwány-a   **chipanda**   **chá**   **chiphadzúwa**
    2SM-pres-7OM-smash-fv   7calabash   7AssocM   7beauty.queen

The ungrammatical sentence involves discontinuity of constituents of a complement of a noun that is itself a complement of a higher noun. This could be handled through appeal to island effects perhaps along the lines of the A-over-A constraint or, possibly, a modified version of the Left Branch Condition of Ross (1967). Alternatively, as noted by Dalrymple (p.c.), the extraction is eliminated by general restrictions on constructions involving non-branching dominance.

Notable from these considerations but not explicitly shown here is the fact that discontinuous constituents in Chichewa and the limits on that do not straightforwardly provide motivation for accounts of information structure grounded in movement rules. Conversely, information structure does not motivate the need for such devices in grammatical theory. This has already been foreshadowed by the observations above concerning constraints on linear precedence. It was noted that the requirement that split NPs preserve the 'base' linear order undermines movement analysis of discontinuous elements (cf. Dahlstrom 1987; Kathol & Rhodes 1999; Mchombo & Morimoto 2004). The movement rule would involve the stipulation that independently moved elements must retain the same ordering relation between the base and derived positions. As noted by Sells (2001), this is a recurrent problem for transformational approaches to phenomena like multiple *wh*-questions and object shift.

6   Conclusion

An issue of recurrent interest in linguistic theory relates to the delimitation of grammatical structure from information structure or discourse. The relation between information structure concepts such as TOPIC and FOCUS with grammatical structure concepts such as SUBJECT and OBJECT, and the realization of these various notions in grammatical structure involves examination of the relation between syntax and pragmatics. Further, the involvement of semantic factors in linearization, and constraints on discontinuity

imposed by, *inter alia*, super-ordinate/hyponym relations, indicate a rather intricate relation between syntax and semantics on the one hand, and between information structure and morphosyntactic organization on the other. This paper has attempted to highlight the intricacies in the relations between information structure and morphosyntactic organization. This is a situation where the relations could be brought into sharp focus in a theory where the informational structures are factored out, with explicit linking principles to capture the connections between the various components.

7   References

Baker, Mark 2003. Agreement, Dislocation, and Partial Configurationality. In Andrew Carnie, Heidi Harley, and MaryAnn Willie (eds.) *Formal Approaches to Function in Grammar: In honor of Eloise Jelinek.* Amsterdam: John Benjamins, 107-134.

Bresnan, Joan 2001. *Lexical-Functional Syntax*. Oxford, UK: Blackwell Publishers.

Bresnan, Joan & Sam Mchombo 1986. Grammatical and anaphoric agreement. *Papers from the Parasession on Pragmatics and Grammatical Theory.* Chicago: Chicago Linguistic Society 22:2, 278-97.

Bresnan, Joan & Sam Mchombo 1987. Topic, pronoun and agreement in Chichewa. *Language 63.4*: 741-82.

Chomsky, Noam 1977. On WH-Movement. In Culicover, Peter, Thomas Wasow & Adrian Akmajian (eds.) *Formal Syntax.* New York: Academic Press, 71-132.

Dahlstrom, Amy 1987. Discontinuous Constituents in Fox. In Kroeber, Paul & Robert E. Moore (eds.) *Native American Languages and Grammatical Typology. Papers from a conference at the University of Chicago.* Bloomington, IN: Indiana University Linguistics Club, 53-73.

Dalrymple, Mary 2001. *Lexical-Functional Grammar* (Syntax and Semantics 34). New York: Academic Press.

De Cat, Cecile 2002. French Dislocation. Doctoral dissertation, University of York.

Deen, Kamil Ud 2004. Let's Build Some Agreement: The status of the Subject Prefix in Nairobi Swahili. Paper presented at the 35th Annual Conference of African Linguistics, Harvard University.

Greenberg, Joseph H. 1963. *The Languages of Africa*. The Hague: Mouton.

Guthrie, Malcolm 1967-71. *Comparative Bantu*. Farnsborough: Gregg International Publishers.

Hale, Ken 1983. Warlpiri and the Grammar of Non-Configurational Languages. *Natural Language & Linguistic Theory* 1: 5-47.

Higginbotham, James 1980. Pronouns and Bound Variables. *Linguistic Inquiry 11*: 679-708.

Jelinek, Eloise 1984. Empty Categories, Case, and Configurationality. *Natural Language & Linguistic Theory 2*: 39-76.

Jelinek, Eloise 1996. Topic and Focus in Navajo Inverse. *Proceedings of the Fifteenth West Coast Conference in Formal Linguistics* (WCCFL 15): 241-255.

Kathol, Andreas and Richard A. Rhodes 1999. Constituency and Linearization of Ojibwe nominals. In Marion Caldecott, Suzanne Gessner, Eun-Sook Kim (eds.) Proceedings of WSCLA 4. *UBC Working Papers in Linguistics II*. Vancouver: Dept. of Linguistics, University of British Columbia, Vancouver, Canada. 75-91.

Kawasha, Boniface 1999a. Relativization and Grammatical Relations in Lunda. Unpublished paper, Dept. of Linguistics, University of Oregon.

Kawasha, Boniface 1999b. Some Aspects of Lunda Grammar. M.A. Thesis, University of Oregon.

Letsholo, Rose 2002. Syntactic Domains in Ikalanga. Doctoral dissertation. University of Michigan.

May, Robert 1985. *Logical Form*. Cambridge, MA: The MIT Press.

Mchombo, Sam 2002. Affixes, Clitics and Bantu morphosyntax. In Mengistu Amberber & Peter Collins (eds.*). Language Universals and Variation. Perspectives on Cognitive Science 3*. Westport, CT: Praeger, 185-210.

Mchombo, Sam 2003. On Discontinuous Constituents in Chichewa. In Zribi-Hertz, Anne & Patrick Sauzet (eds.) *Typologie des langues d'Afrique et universaux de la grammaire* Paris: L. Harmattan, 141-167.

Mchombo, Sam 2004 *The syntax of Chichewa*. Cambridge, UK: Cambridge University Press.

Mchombo, Sam & Yukiko Morimoto 2004. Configuring Topic in the Left Periphery: A Case of Chichewa Split NPs. In Shaer, Benjamin, Frey, Werner & Claudia Maienborn (eds.) *Proceedings of the Dislocated Elements Workshop, ZAS Berlin, November 2003*. ZAS Papers in Linguistics 35.2:347-374.

Mchombo, Sam & Al Mtenje 1983. Noncyclic Grammar. *Linguistic Analysis* 11, 219-236.

Morimoto, Yukiko 2000. Discourse Configurationality in Bantu Morphosyntax. Doctoral dissertation, Stanford University.

Mugane, John 1997. *A Paradigmatic Grammar of Gikuyu*. Stanford, CA: CSLI Publications.

Neale, Stephen 1996. Prolegomena to a "variable-driven" syntax: reflections on logical form and Chomsky's minimalist program. Unpublished paper. University of California, Berkeley.

Ngonyani, Deo 1999. X°-Movement in Kiswahili Relative Clause Verbs. *Linguistic Analysis 29 no. 1-2*: 137-159.

Nolda, Andreas, 2000. Split Topicalization in German. In Sackmann, Robin (ed.) *Studies in Integrational Linguistics*. Amsterdam: Benjamins.

Phillips, Collin 1993. Conditions on Agreement in Yimas. In Bobaljik, Jonathan, D. & Collin Phillips (eds.) 1993. *Papers on Case & Agreement 1. MIT Working Papers in Linguistics Volume 18*. 173-213.

Ross, John R. 1967. Constraints on Variables in Syntax. Doctoral Dissertation, MIT.

Sells, Peter 2001. *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford, CA: CSLI Publications.

mchombo@berkeley.edu