#### JOAN BRESNAN AND SAM A, MCHOMBO

# THE LEXICAL INTEGRITY PRINCIPLE: EVIDENCE FROM BANTU\*

The lexical integrity principle has been called into question by recent work which hypothesizes a syntactic phrasal source for inflected words. Bantu morphology provides a particularly rich empirical domain for this issue because it straddles the boundary between morphology and syntax, inviting syntactic analyses in both the earliest missionary grammars and recent theoretical works in generative grammar (Myers 1987; Baker 1988a,b; Kinyalolo 1991, Carstens 1991). In this study we show that the morphology and syntax of Bantu noun class markers strikingly support the lexical integrity principle, once the morphemic structure of words is factored apart from their prosodic and functional structures.

### 0. INTRODUCTION

How can we tell whether a sequence of morphemes is a word? A fundamental generalization that morphologists have traditionally maintained is' the *lexical integrity principle*, which states that words are built out of different structural elements and by different principles of composition than syntactic phrases. Specifically, the morphological constituents of words are lexical and sublexical categories – stems and affixes – while the syntactic constituents of phrases have words as the minimal, unanalyzable units; and syntactic ordering principles do not apply to morphemic structures. As a result, morphemic order is fixed, even when syntactic word order is free; the directionality of 'headedness' of sublexical structures

<sup>\*</sup> A preliminary version of this paper was given at the Colloquium on Agreement at the Twentieth Annual Conference on African Linguistics at the University of Illinois at Champaign-Urbana on April 21, 1989, and subsequent versions were presented at a UCLA Colloquium in December of 1991 and the Formal Linguistics Society of Mid-America at Northwestern University in May of 1992. We are grateful to Eyamba Bokamba for providing the stimulus and occasion for writing this paper, and to him as well as Alex Alsina, Avery Andrews, Bruce Hayes, Katherine Demuth, Ki-Sun Hong, Larry Hyman, Sharon Inkelas, Jonni Kanerva, Will Leben, K. P. Mohanan, Tara Mohanan, Salikoko Mufwene, Scott Myers, Peter Sells, Whitney Tabor and several *NLLT* reviewers for valuable comments on earlier stages of this work, although we alone are responsible for its shortcomings. We are also grateful to Jonni Kanerva for providing and checking tonal transcriptions of Chicheŵa. Our transcription and glossing conventions for Chicheŵa examples follow Bresnan and Kanerva (1989). This study is based upon work supported in part by the United States National Science Foundation under Grant No. BNS-8919880, Stanford University.

may differ from supralexical structures; and the internal structure of words is opaque to certain syntactic processes.<sup>1</sup>

One source of apparent counterevidence to the lexical integrity principle has come from conflicting criteria used in identifying words. It is clear that the phonological word does not always coincide with the morphological word (Dixon 1977, Selkirk 1984, Nespor and Vogel 1986, Anderson 1984; Booij and Rubach 1984, 1987; Poser 1990; Kanerva 1990; Mohanan forthcoming). There is also evidence that the morphological word does not always coincide with the syntactic word in a certain sense. A single morphological word may incorporate stems or affixes that have the same syntactic functions as phrasal constituents.<sup>2</sup> For example, Bresnan and Mchombo (1987) show that the object marker of the Chichewa verb functions syntactically as an object pronoun rather than a grammatical agreement morpheme, though morphologically it is a verb-stem prefix.<sup>3</sup> Conversely, separate or separable phrasal constituents may constitute a single functional word (Simpson 1983, 1991; Ackerman 1987; Ackerman and Webelhuth 1992, Alsina 1993a,b; Manning 1992; Matsumoto 1992; Butt forthcoming).

These conflicts can be resolved by hypothesizing that the different characterizations of wordhood hold at separate levels of structure. In particular, it has been hypothesized that *the lexical integrity principle* holds of the morphemic structure of words, independently of their prosodic structure or functional structure (Simpson 1983, 1991; Booij 1985; Mohanan forthcoming, Matsumoto 1992)). In the framework of lexical-functional grammar (LFG), which factors apart c(ategorial)-structure and f (unctional)-structure, the lexical integrity principle states that the morphemic structure of words differs from the c-structure of phrases both in constituents and principles of combination. However, morphological and syntactic constituents may have the same grammatical functions at f-structure; indeed, the degree of functional information expressed in word

<sup>&</sup>lt;sup>1</sup> For a range of generative studies in this tradition, see, e.g., Mchombo (1978); Bresnan (1978); Lapointe (1979); Kiparsky (1982); Selkirk (1982); K. P. Mohanan (1982, 1986); Simpson (1983, 1991); Di Sciullo and Williams (1987); Cho and Sells (forthcoming), Sells (1992), T. Mohanan (forthcoming).

<sup>&</sup>lt;sup>2</sup> See Sadock (1980, 1991); Simpson (1983, 1991); Ishikawa (1985); Kanerva (1987); Bresnan and Mchombo (1986, 1987); Demuth and Johnson (1989); Andrews (1990); T. Mohanan (forthcoming); Cho and Sells (forthcoming); Sells (1992); Matsumoto (1992).

 $<sup>^3</sup>$  Evidence for the prefixal status of the object marker can be found in Kanerva (1990) for Chicheŵa and in Myers (1987) for Chishona.

structure or phrase structure is a major parameter of syntactic variation.<sup>4</sup> Convincing evidence for this hypothesis – that lexical integrity is relativized to level of representation – is provided in recent work by T. Mohanan (forthcoming) and Matsumoto (1992).

The theoretical basis of the lexical integrity principle even in this relativized form is called into question by recent work which derives the morphemic structure of words by means of syntactic head movement.<sup>5</sup> In this study we show that the morphology and syntax of Bantu noun class prefixes strikingly support the lexical integrity principle as understood here and raise problems for syntactic theories of word derivation. The Bantu noun class markers are a particularly fruitful domain for investigations of lexical integrity because they straddle the borderlines between syntax and morphology and between inflection and derivation.

### 1. The Problem of the Bantu Noun Class Markers

The Bantu noun class markers mark nominals for number and gender, determining the agreement forms of determiners, modifiers and predicates. The gender classes, though largely formal, can be associated with semantic properties such as animacy, configuration, location, size, plurality or quality.<sup>6</sup> Hence, prefixation of a given noun class marker to a stem can simultaneously determine the syntactic agreement properties of the resulting form (an inflectional characteristic) and change the semantic class of the stem (a derivational characteristic). This mixed inflectional and derivational nature of the markers has long been noted by Bantuists.<sup>7</sup> Thus, Mufwene (1980, p. 248) observes,

... it is one of the roles of noun class prefixes to indicate the change in the lexical meaning of a stem. Not only are the prefixes inflections (showing 'number' for the nouns they delimit), they are also derivational markers, which often play a role similar to that of the deri-

<sup>&</sup>lt;sup>4</sup> This conception of lexical integrity has much in common with Sadock's (1991) autolexical theory, differing primarily in the representation of grammatical functions by means of f-structures rather than categorial structures. The factorization of f-structure from c-structure allows us to distinguish naturally between structure-dependent syntactic principles (e.g., constituent order), which respect lexical integrity, and function-dependent principles (e.g., agreement), which do not.

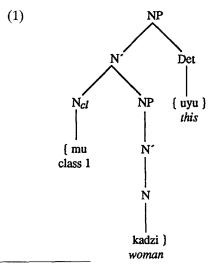
 <sup>&</sup>lt;sup>5</sup> See, e.g., Koopman (1984); Travis (1984); Chomsky (1986); Baker (1988a); Speas (1991).
 <sup>6</sup> The term 'noun class' is often used by Bantuists in place of 'gender'. See Allan (1977) for discussion.

<sup>&</sup>lt;sup>7</sup> See Homburger (1929, p. 44); Doke (1954, p. 51); Welmers (1973, p. 160); Mould (1977); Mchombo (1978); and Mufwene (1980).

vational suffixes -ity, -ment, -er, -hood, -ness, -ation, etc. in English or other Indo-European languages....

In early missionary grammars of Bantu languages, the agglutinative morphemes constituting the noun class markers are sometimes analyzed as separate words and sometimes as prefixes morphologically bound to noun or verb stems. But in virtually the entire tradition of Bantu linguistics, the class markers are analyzed as morphologically bound prefixes, one of the strongest advocates for this position being Doke (1929, 1935).<sup>8</sup>

The standard morphological analysis of the class markers has been challenged in a study of Chishona word structure by Myers (1987). Myers points out that because Doke's criteria for the Bantu word are phonological in character, they establish only that the grammatical elements in question are part of the same phonological word (Doke 1935, p. 22 himself uses the term "phonetic word"), and not that they are parts of a morphological word. A central thesis of generative Prosodic Phonology is that prosodic domains do not necessarily correspond to syntactic constituents (see Inkelas and Zec 1990 and the references cited therein). Myers argues that the Bantu noun class prefixes are syntactically independent and only phonologically bound to their nouns, as illustrated by his analysis of the Chishona word for 'woman', *mukadzi*:



<sup>8</sup> One of Doke's devastating arguments against the syntactic analysis was his pointing out of the inconsistencies in word division in Bantu languages described in French and English missionary grammars. For example, he observes (Doke 1935, p. 11):

Because in English "we are loving" consists of three words, therefore the Englishspeaking missionaries wrote si ya tanda in Zulu and ti no da in Shona. Because in French "nous aimons" consists of two words, therefore the French-speaking missionaries wrote Here the class marker  $N_{cl}$  takes an NP sister in the syntactic structure and forms a phonological (p-) word with the following noun. Thus, *mu kadzi* is analyzed as a sequence of two words syntactically, but as a single p-word for the purposes of phonology. Myers extends this syntactic analysis to all noun class markers, including those which occur with verb stems in synthetic compounds and infinitives/gerunds (class 15).

Myers (1987, p. 12) recognizes that his thesis requires drastic revisions of our understanding of the structure of Bantu and, indeed, of the nature of morphology itself:

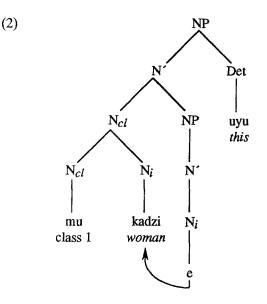
Crucial to this argument will be evidence that the traditional word as defined by Doke and accepted by most subsequent researchers is not in fact a morphological or syntactic constituent at all, but rather a derived phonological domain, i.e. a phonological word. Thus the "word" that has been the focus of almost all previous work on Bantu morphology is not in fact a morphological entity.

This claim raises the central question addressed in this study: How can we tell whether something is a 'morphological entity'? The lexical integrity principle addresses precisely this question.

Kinyalolo (1991) and Carstens (1991) follow Myers (1987) in deriving class-marked nouns and verbs in Kilega and Kiswahili from syntactic configurations in which a nominal head, realized by the classifier, selects an NP or VP complement. But they argue for syntactic head movement of the nominal or verbal stem to the classifier position, forming a syntactic unit of category  $X^0$  by adjunction of the classifier and stem. A generic illustration of the head-movement approach is given in (2). See sections 5, 6 and 7 for more detailed discussion of actual head-movement analyses.

ha randa in Ronga and rea rata in Sotho – despite the fact, that si-ya-tanda contains precisely the same elements as re-a-rata.

See also Bleek (1862); Meinhof (1906); Guthrie (1948); Cole (1955); Welmers (1973, ch. 6).



As Kinyalolo (1991, p. 229) asserts, "... morphological processes of affixation are nothing more than adjunction of an  $X^0$  to a  $Y^{0}$ " in the syntax.

Kinyalolo's and Carstens' analyses, unlike Myers', preserve the view of class-marked noun and verb stems in Bantu as 'morphological entities', in the weak sense of entities dominated by a syntactic category  $X^0$  that also dominates morphologically derived words, but they are in complete violation of the lexical integrity principle defined above. In their analyses, syntactic ordering principles and movement rules apply to morphemic structures, and the syntactic constituents of phrases have not only words, but affixes as the minimal units.

We will show that the evidence from Bantu noun class prefixes supports the standard morphological analysis over the syntactic analysis of most class markers and the lexical integrity principle as understood here over the head-movement theories of word derivation. We begin with a review of five tests for lexical integrity. We then apply these tests to the class markers of nouns in Chicheŵa, which we show to be closely parallel to those in Chishona studied by Myers (1987). The evidence from our lexical integrity tests indicates that prefixal class markers are generated by the morphology side by side with other class markers generated by the syntax. The split cannot be explained by appealing to the difference in syntactic category between NP vs. PP, as proposed by Baker (1988a,b; 1992). And contrary to Kinyalolo (1991) and Carstens (1991), the results cannot be explained by the syntactic head-movement theory of affixation. Next, we show that class-marked verbal constructions evince another split between syntactic and morphological formation which cannot be explained by these syntactic theories of word formation. Finally, we demonstrate the use of the class marker in derivational morphology with Chicheŵa deadjectival nouns, pointing out the consequences for head-movement theories. We conclude that Bantu words are built out of different structural elements and by different principles of composition than syntactic phrases, even though these words and phrases may carry the same information about syntactic functions or relations.

#### 2. FIVE TESTS OF LEXICAL INTEGRITY

We begin with a review of five tests of lexical integrity: extraction, conjoinability, gapping, inbound anaphoric islands and phrasal recursivity.

#### 2.1. Extraction

Constituents of words cannot be extracted by syntactic operations, such as relativization, clefting or topicalization, which leave visible gaps in structure. Thus, the relativization of *American history* in (3) is not possible, whether a gap is left, as in (3b), or a resumptive pronoun, (3c):

- (3)a. They've been [American history] teachers for years.  $\rightarrow$ 
  - b. \*American history, which they've been \_\_\_\_\_ teachers for years, ...
  - c. \*American history, which they've been it teachers for years, ...

The simplest explanation for this constraint is that morphological constituents, being formed in the lexical component, must have lexical content. Empty categories, which receive their content entirely from syntactic context, are thus not available in word formation (assuming base-generation of the movement configuration, as in many syntactic theories).<sup>9</sup>

<sup>&</sup>lt;sup>9</sup> Note that the absence of empty categories in word formation does not rule out 'paradigmatic gaps', in which paradigmatic information may be systematically associated with the absence of morphology (Simpson 1991, pp. 160–2; Andrews 1990). For example, a language may have case inflections for accusative, dative, and genitive, while nominative case is morphologically unmarked. In the context of the case paradigm, the absence of a morpheme carries specific information about the choice of case which need not be modelled by the presence of a null morpheme or empty morphological category. See T. Mohanan (1993) for evidence against a null morpheme for Hindi nominative case. Paradigmatic gaps can also carry information about syntactic functions (see Andrews 1990 for examples), and again there is no violation of the lexical content requirement.

#### 2.2. Conjoinability

Syncategorematic or 'functional' categories, such as conjunctions and complementizers, do not undergo morphological derivation. Consequently, while syntactic categories can be conjoined by syntactic conjunctions, stems and affixes normally cannot (Simpson 1991, p. 52):

- (4)a. Mary outran and outswam Bill.
  - b. \*Mary outran and -swam Bill.
- (5)a. John's joyfulness and cheeriness kept us going.
  - b. \*John's joyful. and cheeriness kept us going.

This phenomenon provides a second test of lexical integrity. The simplest explanation for this constraint is the same as that given for the preceding one: morphological constituents, being formed in the lexical component, must have intrinsic lexical content; complementizers and conjunctions, being syncategorematic categories, lack such content.<sup>10</sup>

There are apparent counterexamples, as illustrated in (6), to the generalization that stems and affixes cannot be coordinated, discussed by Nespor (1985) and Booij (1985), as well as Simpson (1991, pp. 51–55, 57–62):

(6)a. infra e ultrasuoni infra and ultra-sounds

(Nespor 1985, p. 201)

b. Freund oder Feindschaft friendship or hostility (Booij 1985, p. 152)

Nespor's and Booij's work suggests that these are cases of prosodically conditioned ellipsis rather than morphological conjunctions of stems or prefixes. As illustrated in (7), what are conjoined are the complete nouns. The final phonological word  $\omega$  of the first conjunct, which is repeated in the second, undergoes ellipsis after an initial  $\omega$ :

- (7)a.  $(infra)_{\omega}$  e ultrasuoni
  - b. (Freund) $_{\omega}$  oder Feindshaft

This explains why the conjoinable prefixes in Italian are just those which

<sup>&</sup>lt;sup>10</sup> This is not to deny that complementizers and conjunctions may have lexical entries, only that they lack descriptive content of the kind shared by verbs, nouns, adjectives and other lexical categories. The substance of this distinction plays a role in the familiar theoretical distinction between lexical and functional categories (as in Abney 1987, pp. 65ff, for example). Note that the lexical content condition on morphological constituents must not be interpreted to rule out the existence of *affixal* complementizers, conjunctions (see Cho and Sells, forthcoming, for examples in Korean) and determiners (Andrews 1990). This follows if we take morphological constituents, as in Lieber (1983); Inkelas (1989); and elsewhere, to exclude affixes.

belong to separate phonological words from their stems and why the shared suffix in German behaves as a separate phonological word with respect to stress, syllabification and other phonological rules.<sup>11</sup>

#### 2.3. Gapping

A third test of lexical integrity is gapping or ellipsis. As observed by Simpson (1983, pp. 75–77; 1991, pp. 51ff), gapping or ellipsis can apply to syntactic, but not morphological, constituents:

- (8)a. John outran Bill and Mary, Patrick.
  - b. \*John outran Bill and Mary -swam Patrick.

(Simpson 1991, p. 52)

- (9)a. John liked the play and Mary, the movie.
  - b. \*John liked the play, and Mary dis- it. (Simpson 1991, p. 51)

The simplest explanation for this constraint is again the same as for the preceding constraints: the empty or omitted P in (8b) and V in (9b) lack intrinsic lexical content and so cannot be used as morphological constituents. The principles that determine the positioning and expression of the syntactic categories V and P apply in a separate module from the lexical module and to different structural elements, so they have no access to the sublexical V or P in the words *outswam* and *disliked*.<sup>12</sup>

#### 2.4. Inbound Anaphoric Islands

A fourth test of lexical integrity is the inbound anaphoric island constraint (Postal 1969; Ward, Sproat and McKoon 1991): while phrases can contain anaphoric and deictic uses of syntactically independent pronouns, derived words and compounds cannot. This constraint derives from Postal's (1969)

<sup>&</sup>lt;sup>11</sup> The phonological salience of certain English prefixes may also explain their appearance in conjunctions (*pre- and post-World War II*), as suggested by Simpson (1991) and Alsina (1993a), who notes their appearance in other ellipsis contexts (*I'm more interested in prethan I am in post-World War II*). Di Sciullo and Williams (1987) propose that these are ambiguously analyzable as prefixes or independent words.

<sup>&</sup>lt;sup>12</sup> Again, apparent counterexamples may be traceable to prosodically conditioned gapping. See Sadock (1991, p. 98) for an example of "polysynthetic gapping" in West Greenlandic.

observation that derived words are "anaphoric islands". Postal cites contrasting examples like those in (10).<sup>13</sup>

(10) McCarthyite, \*himite (Postal 1969, pp. 213–4)

Note that this constraint applies only to anaphoric and deictic uses of pronouns. Examples of word-internal pronouns, such as *he-man* and *she-bear*, are not excluded because the pronouns are used only for their lexical content of gender and have no indexical function.

There are problems, however, with the formulation of this constraint, which we will illustrate with respect to Bantu. First, morphologically incorporated indexical pronouns occur in Bantu and many other languages (see Bresnan and Mchombo 1987; Kanerva 1987; Demuth and Johnson 1989; Andrews 1990, and the references cited in these). An example is the pronominal object marker prefixed to the verb stem in Chicheŵa (Bresnan and Mchombo 1987, p. 745):

(11) zi-ná-wá-lum-a SB 10-PAST-OB 2-bite-IND They bit them.

Second, in a number of Bantu languages (though not in Chicheŵa), the verb stem together with its pronominal prefix may undergo a derivational morphological process of nominalization. In Chishona, suffixation of *-iro* yields a manner nominalization from a verb stem X, with the derived meaning "way of X-ing", e.g. mu-dy-ir-o "way of eating" (cf. -dy-ir 'eat for"), mu-bvunz-ir-o "way of asking" (cf. -bvunz-ir 'ask for") (Fortune 1985, p. 101).<sup>14</sup> Fortune gives examples with object markers:

- (12) Chishona manner nominalizations (Fortune 1985, p. 101): mu-tí-túm-ir-ó 'way of sending us' (cf. -túm-ír-í 'send for')
  - b. mu-dzí-vák-ir-ó 'way of building them' (cf. -vák-ír-' 'build for')
  - *mu-zví-bát-ir-ó* 'way of holding them things' (cf. -bát-ír-´ 'hold for')

Similarly, in Setswana, deverbal personal nouns formed by prefixing a

(i) \*McCarthy, ites are now puzzled by his, intentions. (Postal 1969, p. 213)

<sup>&</sup>lt;sup>13</sup> Postal (1969) also proposes an 'outbound' anaphoric island constraint – that subparts of derived and compounded words cannot be the *antecedents* of pronouns:

However, we have not used this as a test of lexical integrity because antecedents can be inferentially constructed in appropriate contexts, and the lexical integrity principle provides no theoretical basis for excluding them. In fact, Ward, Sproat and McKoon (1991) argue that outbound anaphora is grammatically permitted and only pragmatically constrained.<sup>14</sup> We are grateful to Scott Myers for providing us with these examples.

class prefix and changing the final vowel of the verb stem may also incorporate the pronominal object prefix (Cole 1955, p. 114).

Third, in Chicheŵa as in other Bantu languages, pronominal stems may be inflected as in (13) or fused with certain noun stems as in (14):

- (13) w-ânga mine (class 1)w-âke his/hers (class 1)
- (14) mnza companion (class 1) mnz-ânga my companion mnz-âke his/her companion

We cannot simply say that these languages lack the constraint Postal observed for English (as suggested by Ward, Sproat and McKoon (1991, p. 450)). In Swahili, for example, the noun *mwana*, literally 'child', can be compounded with a noun N to denote a practitioner of a profession related to N:

 (15) mwana-siasa politician (cf. siasa politics) mwana-sheria lawyer (cf. sheria law) mwana-anga astronaut (cf. anga sky)

We can refer to law, politics or the sky in Swahili by means of the demonstrative pronoun *hii*, but the pronoun itself cannot appear in this compound:

(16) \*mwana-hii Lit.: it-er

In all of the good examples of word-internal pronouns, the pronominal element bears a determinate syntactic function in relation to its associated stem or affix: it is the object of the verb in (11)–(12) and the nominal possessor in (13) and (14). In Bresnan and Mchombo's (1987) analysis, the morphological operation which joins the object marker with the verb stem specifies the content of the syntactic object of the verb as pronominal at the level of functional structure, and the same approach may be extended to the possessive examples.<sup>15</sup> The meaning of these words can thus be derived by semantic composition of syntactic functions in f-structure. In the bad examples of word-internal pronouns in (10) and (16), the pronominal element has no determinate syntactic function in relation to the stem. The meaning cannot be derived by semantic composition of f-

<sup>&</sup>lt;sup>15</sup> Sadock's (1980, p. 314) examples of morphological incorporation of independent nouns in Greenlandic can be analyzed similarly, as proposed by Simpson (1983; 1991; pp. 226– 237). See T. Mohanan (forthcoming) for a similar analysis of Hindi noun incorporation.

structure syntactic functions, but must be a morphological derivative of the stem and affix semantics.<sup>16</sup>

To capture the inbound anaphoric island constraint, we therefore propose that indexical pronouns, though they do have intrinsic lexical content and can appear word-internally, lack the appropriate lexical content to serve as morphological bases for semantic derivatives. This means that the inbound anaphoric island constraint is a valid test of lexical integrity when it is restricted to a subclass of meaning-changing morphological operations that are found in derivation and compounding.<sup>17</sup>

#### 2.5. Phrasal Recursivity

Phrasal recursivity is our last and most controversial test of lexical integrity. On the one hand, word-internal constituents generally differ from word-external phrases in disallowing the arbitrarily deep embedding of syntactic phrasal modifiers:

- (17)a. [A happy]-ness
  - b. \*[AP quite happi]-ness
  - c. \*[AP more happy [than sad]]-ness

This is what would be expected if the morphological constituents of words are restricted to lexical and sublexical categories. On the other hand, evidence that syntactic phrases undergo derivation and compounding has been given for languages such as Afrikaans (Botha 1980, 1981, 1984), English (Spencer 1988, 1991; Lieber 1988, 1992) and Japanese (Shibatani and Kageyama 1988). Lieber (1988, 1992) cites the following examples:

(18) Charles and Di syndrome, employee of the month program, an ate too much headache, I told you so attitude, God is dead theology, stick it in your ear attitude, who's the boss wink

Spencer (1988, 1991, pp. 414–17) suggests that *lexicalization* of a phrase is a precondition for its undergoing word formation, pointing out the contrast in (19a) and one similar to our (19b):

<sup>&</sup>lt;sup>16</sup> See Alsina (1993a) for a discussion of the semantic composition of some morphological derivatives within the general framework assumed here.

<sup>&</sup>lt;sup>17</sup> Simpson (1991, pp. 55–62) shows that Warlpiri pronouns can have 'derivational case' suffixes. Cole (1955, p. 111) cites the use of the class 14 marker in Setswana for derivation of quality nominals from pronouns. These examples may be borderline between a syntactic relational and a semantic derivative analysis.

- (19)a. Baroque flautist (a player of a Baroque flute), wooden flautist(≠ a player of a wooden flute)
  - b. transformational grammarian (a writer of a transformational grammar), partial grammarian (≠ a writer of a partial grammar)

Baroque flautist contrasts with wooden flautist because there is a dictionary entry for Baroque flute and none for wooden flute. Similarly, phrases like transformational grammar are considered to be lexicalized because they have specialized meanings as names of theories or academic subjects that are not predictable from their syntactic structure. In contrast, an unlexicalized phrase, such as partial grammar in She has written a partial grammar of Fijian, lacks a corresponding nominalization: partial grammarian does not mean 'writer of a partial grammar', but 'partial writer of a grammar'. Similar contrasts appear in compounded phrases –

- (20)a. [American history] teacher, ??[recent history] teacher
  - b. [acoustic guitar] player, ??[amplified guitar] player

- although it must be noted that in appropriate contexts all of these phrases would be acceptable.

Clark and Clark (1979), in their study of lexical innovation in denominal verbs such as to porch a newspaper, to Houdini one's way out of a closet, to stiff-upper-lip it, show that there is a gradation of denominal verbs from complete innovations whose interpretation depends on context of use to opaque idioms whose meaning can only be found in a dictionary. They propose principles for deriving the contextually dependent meanings of novel denominal verbs (Clark and Clark 1979, pp. 804–5, 807–8). We see the same gradation in the lexicalization of phrases here. For example, there are morphologically embedded phrases like those in (21):

(21) forget-me-not blue, ne'er-do-well look, know-it-all-ish

These are relatively opaque or well-established lexicalized names for plant life and human types. A number of Lieber's examples in (18) are intermediate cases, drawing on transparent phrases with some currency of usage in advertising or journalism. At the other extreme, the questionable examples we have invented in (20) and (19b) could be used as contextdependent innovations.

The innovative character of these compounded phrases has led Lieber (1988, 1992) to propose general syntactic principles which allow phrasal recursion within lexical categories, in these cases concatenating a syntactic

category XP with a N in violation of the lexical integrity principle. But there are problems with this approach.

First, note that the pronouns in Lieber's examples in (18) are not used indexically. For example, I in the phrase I told you so attitude does not refer to the speaker who characterizes the attitude. The meaning of synthetic compounds of the type [XPN] given by Lieber seems to be 'an N associated with the phrase XP'. This meaning in effect quotes the phrase, encapsulating any indexical expressions it may contain. Such a quasiquotative source for these compounds could also explain the fact that non-English phrases can be innovatively compounded in this way:

(22) a Sturm und Drang romantic, a Heil Hitler skinhead, a mea culpa look, a certain je ne sais quoi quality, his *zóôna!* expression, the ich bin ein Berliner speech

It seems implausible to import German, Latin, French or Chicheŵa syntactic rules into the word formation component of English in order to generate these examples. The same principles that can explain these examples can explain Lieber's examples in (18).

Second, there are many unexplained gaps on Lieber's approach: the relatively interpretable examples given in (18) are syntactically on a par with much worse examples such as those in (23):

(23) ??the Prince of Wales and the woman that he married syndrome,??an ate too much and smoked a post-prandial cigar headache,??who's the manager, proprietor, or CEO wink

Clark and Clark's (1979) work suggests that the solution to these problems is not to import the principles of syntactic formation into word formation as Lieber (1988, 1992) proposes, but to recognize that lexicalization of phrases can be innovative and context-dependent.

In sum, we suggest that true phrasal recursivity is lacking in word structure. Where syntactic phrases appear to undergo morphological derivation, it is by virtue of their being lexicalized. Although lexicalization can be innovative, the nonsyntactic status of lexicalized phrases embedded in word structure can be detected in properties such as lexical gaps, and can be confirmed by the other lexical integrity tests.<sup>18</sup>

We have reviewed five tests of wordhood that can be coherently related to the lexical integrity principle as understood here: extraction, con-

<sup>&</sup>lt;sup>18</sup> These tell-tale properties of lexicalization also appear in the very productive cases of 'postsyntactic compounding' in Japanese discussed by Shibatani and Kageyama (1988) and synthetic compounding in Afrikaans discussed by Botha (1980, 1981, 1984).

joinability, gapping, inbound anaphoric islands, and phrasal recursivity. We turn now to the Bantu phenomena to be investigated.

## 3. Class-Marked Nouns

In this section we apply the lexical integrity tests to class-marked nouns in order to determine whether the class markers are adjoined as part of a syntactic phrase or prefixed morphologically.

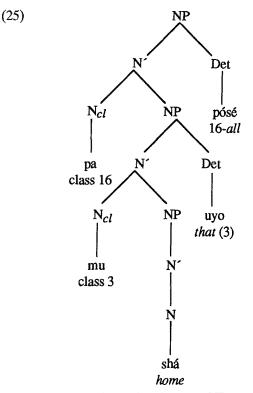
## 3.1. Testing for Phrasal Recursivity

The most striking argument for the syntactic analysis of noun class markers is the existence of alternative concord, in which modifiers may simultaneously show concord with any of several class markers on the same noun. Myers (1987, p. 104) cites examples like the following from the Masvingo Karanga dialect of Chishona:<sup>19</sup>

(24) pa-mu-shá uyo p-ósé p-a-káchén-a 16-3-home that(3) 16-all 16-white

Here the noun *home* has two preceding noun class markers (of classes 16 and 3). The first of its three following modifiers agrees with the inner class 3 marker, and the final two agree with the outer class 16 marker. This phenomenon is neatly accounted for by Myers' analysis of class markers, as shown in (25).

<sup>&</sup>lt;sup>19</sup> The glosses in these and other cited examples follow the originals.



If each class marker takes its own NP complement, which can dominate another class marker, then there are multiple levels of NP in which a determiner or other modifier can appear. The determiner will show concord with the NP which it determines, either the inner or the outer one.

This analysis also entails that the modifiers showing inner concord must precede those showing outer. This fact follows because the concordant modifiers are final in their phrases and the inner class marker NPs are center-embedded in the outer NPs. Thus, the phrasal analysis correctly accounts for the ungrammaticality of examples such as (26), which reverses the concords of the determiners in (24), as in Myers (1987, p. 104):<sup>20</sup>

(26) \*pa-mu-shá apo w-ósé p-a-káchén-a 16-3-home that(16) 3-all 16-white at that whole white home

The class marker 16 in the above examples is one of the locative classes 16, 17 and 18. Alternative concord with the locative markers is widely

<sup>&</sup>lt;sup>20</sup> We have corrected the gloss of *apo* 'that(16)' in this example, which is mistakenly marked as class 3 in Myers' text.

attested in Bantu languages (e.g., Chibemba (Givón 1972), Tshiluba (Stucky 1978; Kuperus and Ilunga 1987), Chishona (Perez 1983) and Kihaya (Trithart 1977)), and also occurs in Chicheŵa (Orr and Scotton 1980; Bresnan and Kanerva 1989). If the syntactic analysis is correct for all class markers, then we should expect to find alternative concord with nonlocative class markers. And indeed Myers (1987, pp. 102, 105) cites two cases from Chishona with nonlocative classes:<sup>21</sup>

- (27) chi-hw-aná u-díki
  7-14-child 14-little
  the ways of small children
  (Fortune 1985, p. 93; cf. hwaná udíki small children)
- (28)a. sá-dunhu uyu {1a-5-*district*}{this (1a)} this district head
  - b. sá-dunhu iri {1a-5-district} {this (5)}
    this head of this district

Finally, Myers (1987, p. 102, n. 30) analyzes in the same way even a single classifier associated with a noun stem, such as *mu-kadzi* '1-woman' in (1). He treats this as "just a special case of this structure in which the noun phrase complement contains nothing but its nominal head. This classless noun phrase cannot include any specifiers or complements since these must bear class features in agreement with the head." Thus, the same syntactic analysis is extended uniformly to all nominal class markers.

Alternative concord is in fact a special case of phrasal recursivity, and it makes a compelling case for the syntactic analysis of the noun class markers by showing their lack of lexical integrity. However, there is reason to question Myers' conclusion that all class markers must have a syntactic status as heads of their own NPs. First, the bare noun stem in Bantu is a bound form which cannot be used uninflected in any syntactic context. Even the citation forms of nouns require a noun class marker. (The English words *Swahili*, *Zulu*, and *Cheŵa* are bare noun stems in the Bantu languages they designate. In these languages they cannot be used without noun class prefixes: *Kiswahili*, *Isizulu*, and *Chicheŵa*.) In the examples above, *-dunhu* 'district', *-aná* 'child', and *-shá* 'home' are such bare noun

<sup>&</sup>lt;sup>21</sup> In (28), Myers shows his parsing of the examples into phonological words.

stems; yet on Myers' analysis, they are full syntactic NPs. Myers' analysis does not explain why in *all* their uses, these NPs must be embedded obligatorily within a higher NP headed by a class marker. Second, there are a number of other class markers in Chishona that appear with nouns already class-marked (the diminutive classes 12 ka- and 13 tu-, the augmentative class 5 *zi*- and the pluralizing class 6 *ma*-), about which Myers is silent. Do these class markers allow alternative concord? It turns out they do not.

In Chicheŵa, as in Chishona, the class markers that appear with already class-marked nouns include the diminutive ka-, ti-, augmentative chi-, zi-, pluralizing ma- and locative pa, ku, mu. But alternative concord is impossible with all except the locative classes, 16, 17 and 18, as shown in the table in (29).

. ,		
outer concord	inner concord	translation
ka-mu-ndá k-ánga 12-3-field 12-my	*ka-mu-ndá w-ánga 12-3-field 3-my	my small field
ti-mi-ndá t-ánga 13-4- <i>field</i> 13-my	*ti-mi-ndá y-ánga 13-4- <i>field</i> 4-my	my small fields
chi-mu-ndá ch-ánga 7-3-field 7-my	*chi-mu-ndá w-ánga 7-3-field 3-my	my large field
zi-mi-ndá z-ánga 8-4- <i>field</i> 8- <i>my</i>	*zi-mi-ndá y-ánga 8-4- <i>field</i> 4- <i>my</i>	my large fields
ma-u-lendó ánga 6-14- <i>journey</i> 6 my	*ma-u-lendó w-ánga 6-14 <i>-journey</i> 14-my	my journeys
pa-nyanjá p-ánga 16-9 <i>lake</i> 16- <i>my</i>	pa-nyanjá y-ánga 16-9 <i>lake</i> 9- <i>my</i>	on my lake
ku-nyanjá kw-ánga 17-9 <i>lake</i> 17- <i>my</i>	ku-nyanjá y-ánga 17-9 <i>lake</i> 9- <i>m</i> y	at my lake
m-nyanjá mw-ánga 18-9 <i>lake</i> 18- <i>my</i>	m-nyanjá y-ánga 18-9 <i>lake</i> 9- <i>my</i>	in my lake

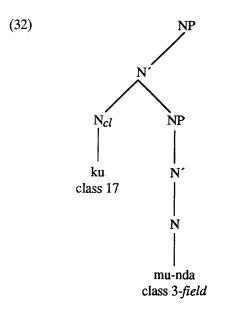
(29)

Simultaneous occurrence of two modifiers differing in concord is also grammatical in Chicheŵa, provided that modifiers showing inner concord precede those showing outer concord, just as in Chishona:

- (30)a. pa mu-dzi p-áthú p-ônse
   16 3-village 16-our 16-all
   at all of our village
  - b. pa mu-dzi w-áthú p-ônse
    16 3-village 3-our 16-all
    at all of our village
  - c. pa mu-dzi w-áthú w-ônse
    16 3-village 3-our 3-all
    at all of our village
  - d. \*pa mu-dzi p-áthú w-ônse
    16 3-village 16-our 3-all
- (31)a. pa mu-dzi p-áthú p-ó-sángálatsa
   16 3-village 16-our 16-ASC INF-please
   at our pleasant village
  - b. pa mu-dzi w-áthú p-ó-sángálatsa
    16 3-village 3-our 16-ASC INF-please
    at our pleasant village
  - c. pa mu-dzi w-áthú w-ó-sángálatsa
     16 3-village 3-our 3-ASC INF-please
     at our pleasant village
  - d. \*pa mu-dzi p-áthú w-ó-sángálatsa
     16 3-village 16-our 3-ASC INF-please

A locative phrase always requires locative concord with verbs and predicate complements, even when its attributive modifiers show nonlocative concord (Bresnan and Kanerva 1989, p. 39).

The failure of recursivity for all but classes 16–18 is explained by the traditional morphological analysis. The nonlocative class markers are prefixed to nouns and noun stems, which lack the recursive structure of syntactic NPs:



We will use the term 'preprefixes' for those prefixes which attach to prefixed nouns, like ka- in (32), rather than to unprefixed stems, like mu-in (32).

The Chishona counterparts of these Chicheŵa preprefixes behave exactly the same way.<sup>22</sup> For example, the class 12 diminutive prefix ka- does not allow concord with the class 5 noun *gudo* 'baboon' to which it is prefixed in (33):<sup>23</sup>

- (33)a. Gudo r-angu ri-no-buda basa.
   5baboon 5-my 5S-PRS-quit 5work
   My baboon is quitting work.
  - b. Ka-gudo k-angu ka-no-buda basa.
    12-5baboon 12-my 12S-PRS-quit 5work
    My little baboon is quitting work.
  - c. \*Ka-gudo r-angu ka-no-buda basa. 12-5baboon 5-my 12S-PRS-quit 5work

<sup>&</sup>lt;sup>22</sup> This observation was communicated to us by an anonymous reviewer of this paper.

 $<sup>^{23}</sup>$  These are the judgments of Haba Musengezi, a first-language speaker of Chishona from Zimbabwe near Harare.

My little baboon is quitting work.

d. \*Ka-gudo k-angu ri-no-buda basa.
12-5baboon 12-my 5S-PRS-quit 5work
My little baboon is quitting work.

Thus, the evidence from phrasal recursivity with class-marked nouns gives us mixed results: the locative class markers support the syntactic analysis, but many other class markers do not. It turns out that other tests for identifying words based on the lexical integrity principle converge on this split.

#### 3.2. Testing for Inbound Anaphoric Islands

The second test which distinguishes true syntactic phrases from derived words is the inbound anaphoric islands phenomenon. Under the syntactic analysis, the inbound anaphoric islands test predicts that anaphoric and deictic uses of pronouns should occur within the phrasal NP complement to a class marker. With the locative class markers, this indeed happens:

- (34)a. mu iyi 18 9this in this (e.g. house)
  - b. pa icho
    16 7*that*on that (e.g. hat)
  - ku ĭwo
    17 6them
    to them (e.g. pumpkins)

But all of the other class markers fail this test. The following examples are representative:

(35)a. \*chi iyi
7 9this
b. \*ka icho
12 7it

c. \*ti ĭwo 13 6*them* 

Once again, this failure is explained by the morphological analysis of these prefixes, given that derived words are inbound anaphoric islands.<sup>24</sup>

#### 3.3. Testing for Extraction

A third test that distinguishes syntactic phrases from derived and compounded words is that morphological constituents of the words cannot be extracted by syntactic operations, such as relativization. However, there are other constraints on extraction that limit the applicability of this test to phrases. In particular, only arguments of verbs can be extracted in Chicheŵa.<sup>25</sup>

To see this, consider first the contrast between the infinitival form of the verb *put*, which allows extraction of its locative object in (36), and the manner nominalization of the same verb, which prohibits extraction of the locative in (37).<sup>26</sup>

- (36)a. Mw-aná w-ánu á-ma-zónd-á
  1-child 1-your 1SB-PRS HAB-hate-IND
  [ku-íká ma-kású m' mǐ-nda].
  15-put 6-hoe 18 4-field
  Your child hates putting hoes in the fields.
  - b. M' mì-nda m-méné mw-aná w-ánu á-ma-zónd-á 18 4-field 18-REL 1-child 1-your 1SB-PRS HAB-hate-IND [ku-ĭka ma-kâsu \_] mu-li mi-kângo. 15-put 6-hoe 18-be 4-lion

In the fields where your child hates to put hoes, there are lions.

In (36a) the infinitive or gerund phrase to put hoes in the fields is the

<sup>&</sup>lt;sup>24</sup> Example (35b) is tonally distinguished from the grammatical ká ícho 'of it (e.g., fool)', which is a syntactic phrase consisting of a prefixed associative marker á followed by an NP. <sup>25</sup> A similar constraint appears to hold in the Kru languages described by Koopman (1984, pp. 21–22). <sup>26</sup> Chicheŵa has object locatives, which permit object agreement on the verb, as well as

<sup>&</sup>lt;sup>26</sup> Chicheŵa has object locatives, which permit object agreement on the verb, as well as oblique locatives, which do not. Thus, we have both  $\dot{a}$ -ma-iká ma-kâsu m'minda 'they put hoes in fields' and  $\dot{a}$ -ma-pézá ma-kâsu m'minda 'they find hoes in fields'. But only the locative of 'put' can be expressed by an object marker:  $\dot{a}$ -ma-mu-iká ma-kâsu 'they put hoes there (e.g., in fields)' constrasts with \* $\dot{a}$ -ma-mu-pézá ma-kâsu 'they find hoes there'. Both types of locative arguments can be extracted by relativization.

object of the verb *hate*. Relativization of the locative in this syntactic phrase is possible, as shown in (36b). In contrast, in (37a) the verb *put* has undergone manner nominalization (Mchombo 1978), and its locative can no longer be extracted, as shown in (37b).

- (37)a. Mw-aná w-ánu á-ma-źond-á
  1-child 1-your 1SB-PRS HAB-hate-IND
  [ka-ik-idwe ká má-kâsu m`mĭ-nda].
  12-put-NOM 12ASC 6-hoe 18 4-field
  Your child hates the way the hoes are put in the fields.
  - b. \*M' mi-ndá m-méné mw-aná w-ánu á-ma-zónd-á
    18 4-field 18-REL 1-child 1-your 1SB-PRS HAB-hate-IND
    [ka-ik-idwe ká má-kâsu ] mu-li mi-kângo.
    12-put-NOM 12ASC 6-hoe 18-be 4-lion

In the fields where this child hates the way hoes are put, there are lions.

The essential difference between the two cases is that the infinitive/gerund phrase is headed by a verb but the manner nominalization is headed by a noun, which is morphologically derived from a verb. Note that the NP *hoes* which is the direct object of the infinitive in (36) must be introduced in the manner nominalization by the adnominal particle  $k\dot{a}$  'of' in (37).<sup>27</sup>

The importance of the verbal argument restriction on extraction is brought out by applied verb constructions, in which NPs which otherwise are not arguments of the verb become so when a suffix is added to the verb stem. For example, in (38a) the NP *arrows* is the object of the instrumental preposition *with*; (38b) shows that it cannot be extracted. In contrast, (39a) shows the same NP as an instrumental object of the applied form of the verb, and (39b) shows that it can now be extracted.<sup>28</sup>

(38)a. Á-ma-ph-á mi-kángó ndí mí-páliro.
 2SB-PRS HAB-kill-IND 4-lion with 4-arrow
 They kill lions with arrows.

 $<sup>^{27}</sup>$  Note that in the manner nominalization, class 12 ka appears with a verb stem and lacks the diminutive meaning of the same marker when used with a noun stem.

<sup>&</sup>lt;sup>28</sup> On the syntactic effects of the applied form of the verb in Chicheŵa, see Baker (1988a,b; 1992); Garrett (1990); Alsina and Mchombo (1990, 1993); Bresnan and Moshi (1990); Alsina (1993a).

204

- b. \*mí-páliro i-méné á-ma-ph-á mi-kángó ndî \_\_\_\_\_\_\_\_
   4-arrow 4-REL 2SB-PRS HAB-kill-IND 4-lion with the arrows which they kill lions with
- (39)a. Á-ma-ph-ér-á mi-kángó mí-páliro.
   2SB-PRS HAB-kill-APPL-IND 4-lion 4-arrow
   They kill lions with arrows.
  - b. mí-páliro i-méné á-ma-ph-ér-á mi-kângo \_\_\_\_\_\_
     4-arrow 4-REL 2SB-PRS HAB-kill-APPL-IND 4-lion
     the arrows which they kill lions with

Similarly, in (40a) the reason phrase because of ignorance is an adjunct of the verb phrase kill lions; (40b) shows that this adjunct cannot be extracted. In contrast, (41a) shows a reason phrase expressed as an argument of the applied form of the verb kill; and (41b) shows that as a verbal argument it can now be extracted.

- (40)a. Á-ma-ph-á mi-kángó chi-fukwá chá u-mbůli.
   2SB-HAB-kill-IND 4-lion 7-reason 7ASC 14-ignorance
   They kill lions because of ignorance.
  - b. \*chi-fukwá chi-méné á-ma-ph-á mí-kângo \_\_\_\_\_
    7-reason 7-REL 2SB-PRS HAB-kill-IND 4-lion
    the reason for which they kill lions
- (41)a. Á-ma-ph-ér-á zi-fukwá z-ámbîri mi-kângo.
   2SB-HAB-kill-APPL-IND 8-reason 8-many 4-lion
   They kill lions for many reasons.
  - b. zi-fukwá zi-méné á-ma-ph-ér-á \_\_mi-kângo
     8-reason 8-REL 2SB-PRS HAB-kill-APPL-IND 4-lion
     the reasons for which they kill lions

Therefore, we do not expect the relativization test to show a contrast between the syntactic and morphological analyses of the class markers of nouns. Extraction of the sister to a preprefix such as ka- (diminutive class 12) would violate the lexical integrity principle, but extraction of the sister to a syntactic noun class marker such as mu (locative class 18) would violate the verbal argument restriction:

(42)a. \*mu-ndá ú-méné mw-aná w-ánú á-ma-zónd-á
4-field 4-REL 1-child 1-your 1SB-PRS HAB-hate-IND
kâ \_\_
12
the little garden which your child hates

b. \*mi-ndá í-méné mw-aná w-ánú á-ma-zónd-á
4-field 4-REL 1-child 1-your 1SB-PRS HAB-hate-IND
ku-íká ma-kású mû \_\_
15-put 6-hoe 18
the fields which your child hates to put hoes in

#### 3.4. Testing for Conjoinability

The fourth test is conjoinability. The syntactic analysis predicts that two NP complements should be conjoinable under a single class marker, just as two NPs are conjoinable under a single preposition.<sup>29</sup> An example of conjoined NP objects of a preposition is given in (43):

(43) Mu-na-chéz-á ndí *II PL/HON SB-REC PST-converse-IND with*[m-phunzitsi kapéná m-sangalatsi]?
1-teacher or 1-entertainer
Did you converse with the teacher or the entertainer?

As expected, the locative classes do allow such conjoined complement NPs:

- (44)a. Mu-ku-pít-á ku *II PL/HON SB-PROG-go-IND* 17
  [m-sika kapéná m-zinda]?
  3-market or 3-city
  Are you going to the market or the city?
  - b. A-na-gw-ér-á m 1SB-REC PST-fall-APPL-IND 18

<sup>&</sup>lt;sup>29</sup> The conjoinability of nominals is constrained by principles of gender resolution (Corbett and Mtenje 1987).

[chi-tsîme kapéná chǐ-gwa]? 7-well or 7-valley Did he fall into the well or the valley?

c. Mu-na-ík-á pa II PL/HON SB-REC PST-put-IND 16 [m-pando kapéná m-tŏndo]? 3-chair or 3-mortar Did you put (it) on the chair or the mortar?

All of the other class markers, in contrast, fail this test. This failure is illustrated by the class 1 prefix in (45):

(45) \*Mu-na-chéz-á ndí *IL PL/HON SB-REC PST-converse-IND with*m- [phunzitsi kapéná sangalatsi]?
l- teacher or entertainer
Di devene server with the teacher or entertainer

Did you converse with the teacher or entertainer?

Similarly, when we apply this test to the preprefixes, such as diminutive ka-, we find that none of them can take scope over conjoined nouns. Thus, example (46) cannot be interpreted as shown, with the prefix ka-applying to both *mpando* and *mtŏndo*:

(46) \*A-na-b-á ka- [m-pando kapéná m-tŏndo]?
1SB-REC PST-steal-IND 12 3-chair or 3-mortar
Did he steal a little chair or a little mortar?

It is also possible to conjoin two locative class markers of the same noun, as in (47a), but preprefixes cannot be conjoined in this way in (47b):

- (47)a. Ndi-na-jámbúl-a zi-thûnzi [mu ndí pa] ISG SB-REC PST-draw-IND 8-picture 18 and 16 ma-dengu.
  6-basket I drew pictures in and on baskets.
  - b. \*Ni-na-gúl-á ka- ndí chi- gálímoto. *I SG SB-REC PST-buy-IND* 12 and 7 5car
    I bought a little and a big car.

This syntactic test correlates exactly with the results of alternative concord, supporting the morphological hypothesis.

#### 3.5. Testing for Gapping

The fifth test is gapping. Under the syntactic analysis of class markers, it should be possible to gap the noun following the class marker. This is indeed possible with locatives:

(48)a. A-nyamăta a-na-vín-á njerero
2-boy 2SB-REC PST-dance-IND 9name of dance
pa bwaló lá mfúmú Kapanga ndí pá (bwaló)
16 5courtyard 5ASC 9chief K. and 16 (5courtyard)
lá mfúmú Kapatuka.
5ASC 9chief K.
The boys danced the nierero dance on Chief Kapanga's court

The boys danced the njerero dance on Chief Kapanga's courtyard and on chief Kapatuka's (courtyard).

b. A-nyamăta a-na-lów-éts-a nkhôsa m'
2-boy 2SB-REC PST-enter-CAUS-IND 10sheep 18
khólá lá mfúmú Kapanga kapéná m' (khólá)
5corral 5ASC 9chief K. or 16 (5corral)
lá mfúmú Kapatuka.
5ASC 9chief K.

The boys drove the sheep either into chief Kapanga's corral or into chief Kapatuka's (corral).

c. Kodí á-na-kánkh-ír-a m-pando ku chi-pinda Q 2SB-REC PST-push-APPL-IND 3-chair 17 7-room chá ána kapéná ku (chi-pinda) chá á-tsíkana? 7ASC 2child or 17 (7-room) 7ASC 2-girl Did they push the chair to the children's room or to the girls (room)?

But none of the other class markers allow this gapping. The following examples are representative:

(49)a. \*A-nyamăta a-na-lámbúl-a chi-bwaló chá 2-boy 2SB-REC PST-sweep-IND 7-5courtyard 7ASC mfúmú Kapanga ndí chí \*(bwaló) chá 9chief K. and 7 (5courtyard) 7ASC mfúmú Kapatuka. 9chief K. The boys swept Chief Kapanga's huge courtyard and chief Kapatuka's huge courtyard.

b. \*Kodí áná awa a-ku-fún-á m-pira w-á
Q 2child 2this 2SB-PROG-want-IND 3-ball 3-ASC
mphira kapéná m-\* (pira) w-á nsanza?
9rubber or 3- (ball) 3-ASC 10rag
Do this children want a rubber ball or a rag ball?

The morphological analysis predicts the failure of gapping in these examples.

These results correlate exactly with the evidence from alternative concord. The same tests that support the syntactic analysis of the locative class markers disconfirm the syntactic analysis of all other noun class markers. Thus, all nonlocative class markers on nouns are prefixes.

#### 4. Source of the Split

Why should the locative class markers be syntactically independent while all the other noun class markers are morphological prefixes?

One possible explanation is that categorial structure is universally projected from principles of grammar which imply that locatives belong to a distinct syntactic category of prepositional phrases (Stowell 1981; Baker 1988a,b; 1992). On this theory the locative class markers fail to behave like the other noun class markers because they are actually prepositions heading a non-nominal category (PP). This theory has the advantage of accounting for why those noun class markers in Chicheŵa that have an independent status in the syntax form a natural semantic class, but it has three serious problems.

The first problem is that it fails to explain why the locative phrases show locative concord at all. If locatives belong to the distinct non-nominal category of PP, why do they allow the full set of concordant nominal modifiers that NPs have?<sup>30</sup> Not only determiners, but interrogatives, relatives, quantifiers, adjectives, possessives, attributive phrases and verbal subject and object concords in Chicheŵa all have a set of locative prefixes exactly parallel in phonology, morphology and syntax to the concordial prefixes of other noun classes.

The second problem is that in Chishona, unlike Chicheŵa, there is evidence that the syntactically independent class markers do *not* form a natural semantic class. As we discussed above, Myers (1987) cites examples of alternative concord with classes 1A (containing mostly humans) and 7 (containing cultural properties) as well as with the locative classes.

The third problem is that in Chicheŵa locative class-marked phrases show the syntactic distribution of NPs.<sup>31</sup> Locatives can be the subjects and objects of verbs (if semantically compatible), and they can be passivized like other subjects and objects as well.<sup>32</sup> Even when we force the syntactic analysis of the locative class markers by adding an alternative concord modifier to a locative phrase, this NP-like behavior continues. Thus, (50a) shows a locative phrase with alternative concord as the subject of a verb (note the locative subject agreement on the verb, which is obligatory). Example (50b) is the passivized version of (50a), showing the same locative phrase as the object of the passive agentive preposition *ndi* 'by'.

- (50)a. Pa mu-dzi w-áthú p-ó-chítítsa chǐ-dwi
  16 3-village 3-our 16-ASC INF-attract 7-interest
  pá-ma-sangaláts-á a-lěndo.
  16SB-PRS HAB-please-IND 2-visitor
  Our interesting village pleases visitors.
  - b. A-lěndo á-ma-sangalats-ídw-á ndí pá mu-dzi 2-visitor 2SB-PRS HAB-please-PASS-IND by 16 3-village

<sup>&</sup>lt;sup>30</sup> Carstens (1991) proposes that the locative classifiers with their nouns are PPs which are complements to a null nominal head, but there are problems with this proposal. First, the null head position can never be instantiated by any real noun; second, true PPs have no gender class and disallow concord; and third, the putative PP can never appear unembedded in an NP (even when the locative is forced into a PP position, it retains its nominal properties, such as having locative concord). See Bresnan (1994a, b).

<sup>&</sup>lt;sup>31</sup> This point is argued in detail in Bresnan (1991, 1994a, b). Baker (1992) offers a revised analysis of locatives in three Bantu languages, but his proposals for Chicheŵa and Sesotho are inconsistent with the evidence given here, and in Bresnan (1991, 1993).

<sup>&</sup>lt;sup>32</sup> This has been frequently observed in Bantu languages. See, for example, Dalgish (1976a,b); Dalgish and Sheintuch (1976); Stucky (1976, 1978); Hodges and Stucky (1979); Trithart (1976); Alsina (1993a).

w-áthú p-ó-chítítsa chǐ-dwi.
3-our 16-ASC INF-attract 7-interest
Visitors are pleased by our interesting village.

Likewise, the object locatives share the syntactic properties of other objects, including adjacency to the verb, object agreement and passivizability (Baker 1988a,b; Alsina and Mchombo 1990, 1993). And this NP-like behavior persists when the locative phrase has alternative concord modifiers. Thus, in (51a) a locative phrase with alternative concord appears with locative object agreement on the verb. In (51b) the same phrase is the subject of the passive verb. In (51c) another such locative phrase appears as an applied object (with either a beneficiary or locative role); with the beneficiary interpretation, the locative phrase must be adjacent to the verb, exactly as with applied NPs of other classes (Alsina and Mchombo 1990, 1993).<sup>33</sup>

- (51)a. A-lěndo á-ma-pa-kônd-a pa mu-dzi
  2-visitor 2SB-PRS HAB-16OB-love-IND 16 3-village
  w-áthú p-ó-chítítsa chǐ-dwi.
  3-our 16-ASC INF-attract 7-interest
  Visitors love it, our interesting village.
  - b. Pa mu-dzi w-áthú p-ó-chítítsa chǐ-dwi 16 3-village 3-our 16-ASC INF-attract 7-interest pá-ma-kond-édw-á ndí á-lěndo. 16SB-PRS HAB-love-PASS-IND by 2-visitor Our interesting village is loved by visitors.
  - c. A-ku-lúk-ír-a pa m-chenga w-áthú 2SB-PROG-weave-APPL-IND 16 3-sand 3-our p-ó-sángálatsa u-konde.
     16-ASC INF-please 15-net

They are weaving a net for/on our pleasing beach.

This behavior of locative phrases is completely inexplicable if they are analyzed as PPs, a non-nominal or 'case-resistant' category (Stowell 1981;

<sup>&</sup>lt;sup>33</sup> Example (80c) is judged somewhat questionable when the locative modifier 'pleasing' is omitted from the locative phrase 'our beach', perhaps because of a processing problem of local ambiguity resolution of the complements following the applied verb.

Baker 1988a,b). For example, the instrumental PP in Chicheŵa in (52a) never allows object marking, as in (52b), can never be passivized, as in (52c), and cannot be adjacent to the verb without a marked stylistic effect of focus:

- (52)a. A-na-gúmúl-a nyǔmba ndí mí-kôndo.
   2SB-REC PST-demolish-IND 9house with 3-spear
   They demolished the house with a spear.
  - b. \*A-na-ú-gúmul-á nyúmba ndí m·kôndo. 2SB-REC PST-3OB-demolish-IND 9house with 3-spear
  - c. \*(Ndí) m-kôndo u-na-gúmúl-idw-á nyúmba. (with) 3-spear 3SB-REC PST-demolish-PASS-IND 9house

Myers (1987, p. 85) also argues in favor of the nominal analysis of locatives in Chishona, pointing out that like other NPs, which (in GB terms) need Case, locative-marked nouns cannot be complements to nouns without insertion of the associative ('genitive') marker. The same is true in Chicheŵa:

- (53)a. ku mu-dzi kw-âthu 18 3-*village* 18-*our* at our village
  - A-na-fík-á ku mu-dzi kw-âthu.
     2S-REC PST-arrive-IND 18 3-village 18-our They arrived at our village.
  - c. \*mw-aná kú mu-dzi kw-âthu
    1-child 18 3-village 18-our
    ≠ a child at our village
  - d. mw-aná w-á kú mu-dzi kw-âthu
    1-child 1-ASC 18 3-village 18-our
    a child from our village

As we see in (53d), an associative marker must intervene between the head noun and the locative phrase modifier. The associative marker itself has been analyzed by Myers (1987) as a preposition that allows variable gender inflection. In contrast to these locatives, true PPs in Chicheŵa

cannot appear as adnominal modifiers at all, with or without an associative marker:<sup>34</sup>

(54)a. \*mw-ăna ndí m-peni1-child with 3-knifea child with a knife

b. \*mw-aná w-á ndí m-peni
 1-child 1-ASC with 3-knife
 a child with a knife

Further evidence for the NP status of the locative in Bantu can be found in Bresnan and Kanerva (1989) and Bresnan (1991, 1994a, b) for Chicheŵa, and in Demuth and Mmusi (1992) for Setswana.

For these reasons we reject the prepositional phrase analysis of the locative class markers.

There is an alternative explanation that can account for all of the evidence we have considered. Greenberg (1977, 1978) hypothesizes that the class markers in Niger-Congo have evolved historically from syntactic elements of NPs, namely classifying determiners or articles, which became morphologically bound as prefixes or suffixes. If this is so, the class marker prefixes would fit in with the widely observed historical evolution of agreement markers from pronominal or determiner elements (Lehmann 1982, 1988; Givón 1976; Bresnan and Mchombo 1986, 1987; Heine and Claudi 1986). A syntactic word first becomes phonologically reduced and bound to an adjacent constituent and then becomes morphologically bound as an affix. Given Greenberg's hypothesis, the simplest explanation for the syntactic behavior of the locative class markers is simply that this process of historical change was not completed for all the class markers of proto-Bantu: most became prefixes; a few, including the locatives, retained their syntactically independent status as nominal constituents.

The assumption that the locative class markers are syntactic elements of NPs and not PPs solves all three problems that beset the prepositional theory. First, locative phrases show locative concord because they are NPs, and all Bantu NPs have noun class concord. Second, other class markers than locatives may show alternative concord, as in the Chishona examples cited by Myers, because they, too, are incompletely morphologized, either in the current state of the grammar (if the alternation is

 $<sup>^{34}</sup>$  Example (54a) is grammatical under the reading 'a child and a knife', where *ndí* is the conjunction 'and' rather than the preposition 'with'.

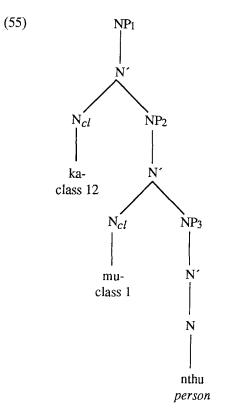
syntactically productive) or in a historically earlier state (if the alternation is fossilized). And third, locative phrases can appear in the syntactic positions of NPs (as subjects, objects, and prepositional objects) just because they *are* NPs.

Thus, the split we have observed between the syntactic and morphological class markers cannot be reduced to the categorial difference between NP and PP. Instead, the prefixal class markers are generated by the morphology side by side with other class markers generated by the syntax. Both types are nominal, and both have the same function in the Bantu concord system.

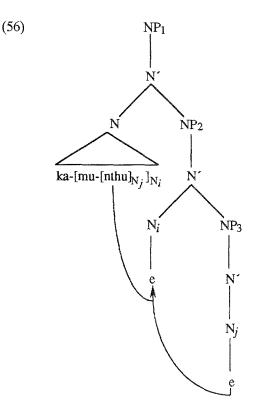
The fact that agreement is marked both syntactically and morphologically in Chicheŵa does not violate the lexical integrity principle assumed here. Recall that lexical integrity holds of the morphemic structure of words and not their prosodic and functional structures. By factoring apart the syntactic levels of f-structure and c-structure, we can distinguish naturally between structure-dependent syntactic principles (e.g., constituent order), which respect lexical integrity, and function-dependent syntactic principles (e.g., agreement), which do not.

#### 5. The Head Movement Analysis

Let us now turn to the theory that verbs and nouns receive their inflections by syntactic movement of lexical heads to the heads of lexical or functional projections in extended X' theory, which has been applied by Kinyalolo (1991) and Carstens (1991) to the analysis of Bantu noun class prefixes. On this theory, the syntactic analysis of the Bantu noun class markers holds, not at surface structure, but at an underlying level of structure. Structures like (55), which we have shown do *not* characterize the constituency of Chicheŵa words, are not the final level of analysis.



From (55), the morphological word could be derived as shown in (56):



First, the noun stem  $N_j$  is adjoined to the inner classifier, and then the resulting complex,  $N_i$ , is adjoined to the outer classifier. Such an analysis is proposed by Carstens (1991, p. 37) for Kiswahili.<sup>35</sup> Superficially, this

<sup>&</sup>lt;sup>35</sup> Carstens' proposal differs in details that are not essential to our main point here. She hypothesizes a functional projection for number (#P) in the phrase structure of the NP, to which the noun stem raises. While number is a syntactic projection, gender is a lexical feature of noun stems. She thus analyzes the class markers as gender-selected allomorphs of the number morpheme, which do not themselves determine the gender of the stem. This analysis is problematic for cases where the stems do not appear to carry the appropriate gender information, as with the preprefixes, locatives, and infinitives. For each preprefix, Carstens therefore hypothesizes a corresponding null N to carry the appropriate gender information. (For instance, the diminutive ka- marks the class 12 gender of a null N which heads NP1 within (55).) For those Bantu languages in which the preprefixes appear directly as bare stem prefixes or, alternatively, as prefixes and preprefixes (e.g., Kikuyu (Carstens 1991, p. 39)), she postulates another set of null derivational morphemes (suffixes) to carry the inherent gender features. But the latter set of null morphemes cannot be identified with the former (pre-stem) set - a clear loss of generalization. These variations are explicable if we make the traditional assumption that it is the prefix ka- that carries the class 12 gender information and diminutive meaning. This remains constant whether ka- is prefixed at the stem or word level. In Southern Bantu languages such as Isizulu, in which proto-Bantu ka-

proposal might seem to be watertight: where there is evidence of lexical integrity, head movement applies to join the stem and classifier into a single word; where there is evidence against lexical integrity, movement does not apply. On this view the entire issue of lexical integrity appears orthogonal to the syntactic theory. On a less superficial view, however, the evidence tells a different story.

According to the syntactic theory, the movement of each stem leaves behind a phrasal structure – NP<sub>2</sub> in (56) – that could contain specifiers, complements and modifiers. The presence of such modifiers would provide no barrier to movement. Indeed, the possible 'stranding' of nonhead material by head movement is built into the theoretical design of the syntactic  $X^0$  movement theory and is essential to obtaining its major empirical results in syntax (such as explaining word order variations in Germanic and Kru (Haider and Prinzhorn 1986; Koopman 1984) as well as in noun incorporation (Baker 1988a)).

We have already seen, however, that such inner modifiers cannot appear in (56). These are the cases of alternative concord that are ruled out for all classifiers but the locatives in Chicheŵa and a few others in Chishona (see examples (27), (28) and (29) above). The failure of alternative concord is explained by the lexical integrity principle but is clear evidence against the syntactic analysis in (56).

One might think of modifying the syntactic analysis by proposing that the classifier selects for NP<sub>2</sub>, as in (56), but that all the specifiers and modifiers come in at a higher level, such as DP (Abney 1987) or some other functional projection. But what then of the specifiers and modifiers at the NP<sub>2</sub> and NP<sub>3</sub> levels assumed in (56)? First, if there are such structures, they will show concord with their head – as all nominal specifiers and modifiers do in Chicheŵa, not only determiners, possessives and relatives, but quantifiers, numerals, adjectives and participles. The presence of any of these at the inner NP levels will give rise to alternative concord, which is ungrammatical. Second, if there are no such structures, then there is no evidence for phrasal NP structure at all: the classifier simply selects for a nominal stem – a morphological unit, not a syntactic one.

The evidence from coordination is equally inexplicable under the syntac-

has been lost, alternative diminutive morphology is used. For example, Isizulu uses the morpheme *ana* 'child' as a diminutive derivational suffix without changing the gender of the stem to which it attaches: *umfana* 'boy (class 1)', *umfanyana* 'little boy (class 1)'; into 'thing (class 9)', *intwana* 'little thing (class 9)'; *amazwi* 'words (class 6)', *amazwana* 'few words (class 6)' (Doke 1988, p. 73). Carstens' analyses of locative (class 16–18) and infinitive (class 15) gender are also problematic (see n. 48 and n. 30).

tic analysis.  $X^0$  categories can be conjoined in Chicheŵa, just as they can in English.<sup>36</sup> Given this fact, nothing prevents  $X^0$  movement from applying to a conjunctive  $X^0$  category in the position of N<sub>1</sub> in (56) and adjoining it to a higher classifier position, yielding ungrammatical examples such as (57):<sup>37</sup>

(57) \*a-[nyamátá ndí tsíkána] a-nâyi
 2-boy and girl 2-four
 four boys and girls, i.e. a group of four consisting of boys and girls

Kinyalolo (1991, p. 231) suggests that by assuming that the class prefix is inflectional and is a spell-out of the number feature on the head N, as proposed in Carstens (1991), the occurrence of the class prefixes on each conjunct follows from across-the-board feature-copying. But this proposal offers no explanation for why other classifiers in the same inflectional concord system *can* be stranded by conjunction and gapping. Merely labelling a morpheme 'inflectional' does not suffice to explain its morphological properties on the head-movement theory. The theory must be supplemented with further principles to account for why syntactically created adjunction structures behave like morphological words – behavior that is predicted by the lexical integrity principle.

As for inbound anaphoric islandhood, pronouns do not undergo a class of word derivation processes, but no principle prevents their occurrence in the syntactic configurations of the movement theories. Thus, the fact that the prefixing Bantu classifiers create inbound anaphoric islands is not a consequence of the movement theory. By the same token, lexically empty syntactic categories do not undergo word formation, but they can occur in the syntactic configurations of the movement theory, leaving the gapping phenomena unexplained. Stipulations must be added to the syntactic theory prohibiting the adjunction of empty syntactic categories to prefixes. This property is a consequence of the lexical theory of word formation.

These facts tell us that something is amiss with the fundamental premise of the head-movement theory, that the stems receive their prefixes through movement operations over a syntactic domain. The theory fails to exclude from this theoretical domain just those syntactic elements – phrasal speci-

<sup>&</sup>lt;sup>36</sup> See (47a) for an example.

<sup>&</sup>lt;sup>37</sup> Carstens' (1991) suggestion that conjoined phrases are barriers to head movement is inapplicable here because the  $X^0$  movement of the conjoined  $X^0$  categories does not cross such a barrier.

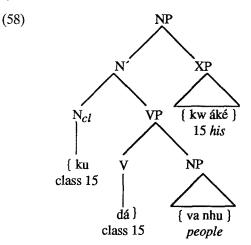
fiers and modifiers, conjoined categories, empty phrasal and lexical categories and pronouns – that are absent from the lexical domain on which morphological principles operate under the lexical integrity principle. Dividing the principles of grammar into a morphological and a syntactic module is not sufficient to recapture the lexical integrity principle, so long as both modules operate over the same domain of structures. What is required is a deeper modularization of the structural elements themselves.

## 6. CLASS-MARKED VERBAL CONSTRUCTIONS

We have established that class-marked nouns are generated both by the syntax and by the morphology. Though the two types have the same function in the concord system, the evidence from lexical integrity tests indicates clear differences in their structural formation. We will now show that nominalized verbal constructions also split across the morphology-syntax divide in their structural formation.

In Chicheŵa, Chishona, and Bantu more generally, noun class markers are used to nominalize verbs or verbal constructions. The class 15 marker ku- derives infinitives/gerunds from verb stems, and other noun class prefixes appear to nominalize VPs, forming what Myers (1987) terms synthetic compounds.

Infinitive phrases show both the internal structure of VPs and the external structure and distribution of NPs, and this duality is captured by Myers' analysis for the Chishona phrase 'his loving people' shown in (58):<sup>38</sup>



<sup>&</sup>lt;sup>38</sup> Myers analyzes all genitive phrases as PPs; without prejudicing the issue, we leave the category of this and other concordant NP modifiers unspecified.

The VP-like properties include the fact that the verb can take any kind of verb complement, such as a direct object. The NP-like properties include the appearance of the concordant possessive, which ordinarily modify nouns, and the fact that these infinitives can occupy NP positions in the syntax. Both of these properties are illustrated in the Chishona example given in (59), cited by Myers (1987, p. 96):

(59) Ku-kora kw-áké ku-kúrú kú-no-shámisá.
15-fatten 15-his 15-great 15-HAB-amaze
His great stoutness is astonishing (cf. Fortune 1955, p. 99)

Here the infinitive/gerund kukora 'to fatten' is modified by a possessor and an adjective, both showing class 15 concord with it, and the entire phrase is the subject of the verb kúnoshámisá 'be amazing', which shows class 15 subject agreement with it.

Myers' syntactic analysis provides an elegant account of this duality. Whereas he has argued, and our lexical integrity tests have confirmed, that at least some class markers are nominal elements syntactically adjoined to NPs, the class 15 marker would be a nominal element syntactically adjoined to a VP. This would neatly explain the duality of these constructions.

Chicheŵa infinitives/gerunds exhibit the same duality in verbal and nominal properties that Myers finds in Chishona.<sup>39</sup> Virtually any verb phrase in Chicheŵa has a corresponding infinitive/gerund phrase form (class 15), which can contain direct objects and other verbal complement types. At the same time, these class 15 phrases can trigger object agree-

(37) ku-li-nyesha mvua jana
 17-past-rain 9 rain yesterday
 It rained yesterday.

<sup>&</sup>lt;sup>39</sup> It appears that Kiswahili gerunds share this duality, but Carstens' (1991) analysis of Kiswahili *poss-ing* gerunds treats them as purely verbal. She admits that this poses a problem for explaining gender concord:

Under the account that I have proposed, it is surprising that agreement with a *poss-ing* type ku- phrase is possible, since only [+N, -V] items have gender specifications (see 2.5.1). It happens that the 'there'-type of expletive agreement in Kiswahili is homophonous Class 17 ku- agreement, [sic] however:

I propose that so-called Class 15 agreement is really expletive, Class 17 agreement, and that the null nominalizer of *ing-of* constructions is of this class also. If this is true, Class 15 does not exist in Kiswahili at all.

This account is inadequate in general (n. 48).

ment (60a) and subject agreement (60b) with the verb,<sup>40</sup> and can be modified by concordant nominal possessors (60c), determiners (60d) and attributives (60e):

(60)a. Ku-dyá onga mw-aná w-ánú 15-eat 14gunpowder 1-child 1-your a-ku-kú-kónd-a. 1SB-PROG-15OB-love-IND Eating gunpowder, your child loves it.

220

- b. Ku-dyá onga kú-ma-kond-êdw-a
  15-eat 14gunpowder 15SB-PRS HAB-love-PASS-IND
  ndí mw-áná w-ánu.
  by 1-child 1-your
  Eating gunpowder is loved by your child.
- c. A-ku-kónd-á ku-píndá njingá kw-áko.
   2SB-PROG-like-IND 15-bend 10bicycle 15-your
   S/he likes your bending bicycles.
- A-ku-kónd-á ku-píndá njingá uku.
   2SB-PROG-like-IND 15-bend 10bicycle 15this
   S/he likes this bending bicycles.
- e. A-ku-kónd-á ku-píndá njingá 2SB-PROG-like-IND 15-bend 10bicycle k-ó-séketsa.
  15-ASC INF-make laugh S/he likes the comical bicycle bending

Myers reinforces this argument for a syntactic analysis by pointing out that other class markers can take VP or even sentential complements in the formation of synthetic compounds. He provides the following Chishona examples (Myers 1987, pp. 99–100, 108):

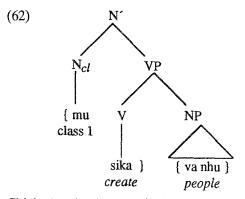
(61)a. Mu-sika vanhu 1-create people

<sup>&</sup>lt;sup>40</sup> See Bresnan and Mchombo (1987) on the syntax of object marking in Chicheŵa.

creator of people, i.e. God (cf. ku-sika vanhu *to create people*)

- b. mu-baya mhondoro
  3-pierce lion
  lion piercer, i.e. acacia thorn tree
  (cf. ku-baya mhondoro to pierce a lion)
- c. chi-ma-ndí-ón-ér-e-pí
   7-2ndPL-1stSG-see-APPL-FV-where
   pygmy
   (cf. mandíónérepí where did you see me?)

He argues that the string of words following the class marker in such examples can contain anything found in a syntactic phrase: a verb and object, as in (61a,b), or even the components of a sentence, as in (61c). He concludes that the class markers in these examples take syntactic phrasal complements of the type VP or S', although they may form a phonological word only with an initial part of that phrase. The structure of (61a) would be as shown in (62):



Chicheŵa also has nominal compounds formed from syntactic phrases, such as VPs. A number of these are cited by Mchombo (1978, p. 284), from which the table in (63) is adapted.

(63)

tola nkhâni	$\rightarrow$	mtolankhâni
pick news		journalist
sunga chǔma	 →	msungachŭma
keep wealth		treasurer
keep wealin		
ponda matîki	$\rightarrow$	mpondamatîki
step on tickeys		millionaire
(threepenny bits)		
kweza lamba	$\rightarrow$	mkwezalamba
raise up belt		belt-tightener
pinga njira	$\rightarrow$	mpinganjira
block path/way		pathblocker
biock puill way		
lowa m'mǎlo	$\rightarrow$	mlowammălo
enter in place		substitute
-		(pronoun)
kholo no mučlo		mkhalapamwăla
khala pa mwăla	$\rightarrow$	
sit on stone		stone-sitter
khala pa kăti	$\rightarrow$	mkhalapakăti
sit in middle		adjective
1 1 .		
pha dzûwa	$\rightarrow$	chiphadzûwa
kill sun		beauty-queen
dya onga	$\rightarrow$	chidyaonga
eat gunpowder		gunpowder-eater
swa bumbu	$\rightarrow$	chiswabumbu
break vulva		vulva-breaker

Like Chishona, Chicheŵa also has a term for the San people that is cognate with the Chishona expression (61c):

 (64) A-mw-a-ndí-ón-er-á pa-ti.
 2-II PL/HON SB-PRF-ISG OB-see-APPL-IND 16-Q the where-did-you-see-me-from people, i.e. San

It is worth pointing out one obvious difference between the compounds in (61)–(64) and the infinitives/gerunds in (58)–(60). The compounds have a special meaning which is distinct from that of infinitives/gerunds and

from the referring NPs that derive from infinitives/gerunds. The compounds are all either lexicalized terms or sobriquets. For example, the Chicheŵa word for 'millionaire', *mpondamatîki*, is clearly lexicalized. The component word *matîki* 'threepence coins' fell into disuse following the abandonment of colonial English currency. Similarly, *chipondamthengo* 'a consultant fee' is a lexicalized term that is only distantly related to the original meaning designating a payment for a practitioner of traditional medicine to go into the bush to seek herbal medicines. On the other hand, words like *chidyaonga* 'gunpowder eater' and *chiswabumbu* 'vulvabreaker' are used as humorous appellations for a fearless individual and a large penis. Although elaborate compounds can be formed creatively, they still are regarded as descriptive names or epithets, rather than ordinary descriptions.

When we apply our lexical integrity tests, the pattern of evidence indicates that there are structural differences as well between the two types of nominalized VPs. First, the two types differ with respect to the extraction test. In (65a) the infinitive/gerund phrase *kudyá onga* 'to eat gunpowder' is the object of the verb *zonda* 'hate'. Relativization of the object in this syntactic phrase is possible, as shown in (65b).

- (65)a. Mw-aná w-ánu á-ma-zónd-á
  1-child 1-your 1SB-PRS HAB-hate-IND
  ku-dyá onga.
  15-eat 14gunpowder
  Your child hates eating gunpowder.
  - b. Onga u-méné mw-aná w-ánu 14gunpowder 14-REL 1-child 1-your á-ma-zónd-á kŭ-dya u-li ndí nsabwe. 1SB-PRS HAB-hate-IND 15-eat 14-be with 10louse The gunpowder which your child hates to eat has lice.

In a synthetic compound, in contrast, extraction of the object of the verb is impossible. For example, the nominal *chidyaonga* 'gunpowder eater', which is the object of the verb in (66a), completely disallows relativization of *gunpowder*, as shown in (66b):

(66)a. Mw-aná w-ánú á-ma-zónd-á chi-dya-onga.
1-child 1-your 1SB-PRS HAB-hate-IND 7-eat-14gunpowder
Your child hates the gunpowder eater.

#### JOAN BRESNAN AND SAM A. MCHOMBO

b. \*Onga u-méné mw-aná w-ánú 14gunpowder 14-REL 1-child 1-your chí-dya \_ u-li ndí nsabwe. á-ma-zónd-á 1SB-PRS HAB-hate-IND 7-eat 14-be with 10louse The gunpowder which your child hates the eater of has lice.

The only difference between the two examples is the choice of class marker. If both class markers are nominal heads taking VP sisters, as Myers (1987) proposes, it is hard to see what explains this difference in extractibility. In both examples gunpowder is an argument of the verb eat, so the verbal argument restriction on extraction in Chicheŵa should be satisfied in both.41,42

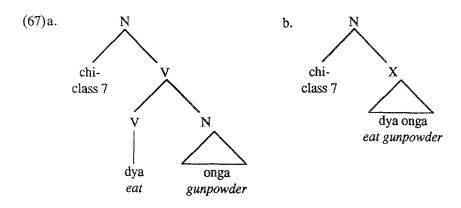
The contrast can be explained by an alternative analysis. Mchombo (1978) and Sproat (1985) argue for a morphological analysis of the Chicheŵa compounds, as illustrated in (67) for gunpowder eater:

<sup>&</sup>lt;sup>41</sup> Kinyalolo (1991, p. 233) proposes that a "Condition on Chain Links" is violated by extraction from synthetic compounds because they are dominated by NPs, and adjunction to NP is not permitted for these extractions. However, these assumptions provide no explanation for the contrast with infinitives/gerunds which are also dominated by NPs, as argued in further detail below.

<sup>&</sup>lt;sup>42</sup> Carstens (1991, pp. 63-65) presents examples of relativization from synthetic compounds in Kiswahili using a genitive resumptive pronoun, as in (i):

kisima ni-na-ye-m-penda mchimba wake (i) hiki ni 7this COP 7well 1S-PRES-7RA-10A-like 1digger This is the well whose digger I like.

But our informants consider these examples clearly ungrammatical. Note that the prefixed verb m-chimba 'class 1-dig' of the synthetic compound normally does not take a genitive complement of any kind, pronominal or otherwise, and it does not occur in isolation as a noun. However, the related nominization of m-chimba-ji 'digger' does take a genitive complement like other deverbal nouns. Replacing \*mchimba in (i) with mchim-baji 'digger' restores the grammaticality of the sentence. All of Carstens' examples of relativization from synthetic compounds depend on this misanalysis. Our analysis correctly predicts their ungrammaticality: as morphological words they disallow extraction by lexical integrity, and the resumptive pronoun strategy of relativization is ruled out by inbound anaphoric islandhood. Only by interpreting them as -ji nominals can one make sense of the examples, and this, we infer, is what Carstens' informants have done.



Sproat (1985) assumes that the phrasal elements are compounded into a V, as in (67a), but we can leave open the exact internal analysis of the lexicalized phrase, dominated by X in (67b). Both analyses in (67) differ crucially from Myers' in that the class marker is a prefix to a compounded or lexicalized phrase, rather than the head of a N' projection taking a VP complement. Such a morphological analysis would explain why the synthetic compounds differ from the infinitives/gerunds with respect to extractibility.

A second difference between the infinitives/gerunds and the compounds shows up with the inbound anaphoric island test. Recall that English phrases like *eater of them, climber of this* contrast with compounds like *\*them-eater* (cf. *man-eater*), *\*this-climber* (cf. *rock-climber*). By this test, the Chicheŵa synthetic compounds are clearly morphologically derived words, in contrast to the infinitives/gerunds, as shown in the table in (68).

1	۷	0	i)
1	O	$\circ$	1
٦.	~	~	1

kondá máûngu	$\rightarrow$	mkondamaûngu
love pumpkins		pumpkin-lover
kondá íwo	$\rightarrow$	*mkondaĭwo
love them (e.g. pumpkins)		them-lover
pinga njira	$\rightarrow$	mpinganjira
block path		pathblocker
pinga iyi	$\rightarrow$	*mpingaiyi
block this (e.g. path)		this-blocker

Prefixed object pronouns are equally impossible in these nominalizations: \**m-wa-konda* 'them-lover', \**m-i-pinga* 'it-blocker'. A prefixed object pronoun does occur in example (64), but is interpreted nonindexically.<sup>43</sup> In contrast, in infinitive/gerund phrases, any NP can be replaced by an anaphoric or demonstrative pronoun, as illustrated in (69):

- (69)a. A-ku-kónd-á ku-píndá izi.
  2SB-PROG-like-IND 15-bend 10DEM PL
  S/he likes bending these.
  - b. Mw-aná w-ánu á-ma-zónd-á ku-dyá uwu.
     1-child 1-your 1SB-PRS HAB-hate-IND 15-eat 14DEM
     Your child hates eating this.

This difference, too, is unexplained by the syntactic analysis of synthetic compounds given above, but it follows from the morphological analysis and the lexical integrity principle assumed here.

A third difference between the two types of nominalized VPs is phrasal recursivity. Mchombo (1978, pp. 284ff) has observed that the compounds fail to show the productive recursivity of true phrasal constituents and therefore should not be analyzed as containing syntactic VPs. For example, we note that alongside *mkhalapamwăla* 'a rock-sitter', there is no *\*mkhalapamwalá wóyêra* 'a sitter on a white rock'; similarly, there is no *\*mpinganjira zókhôta* 'a blocker of crooked paths'. To take another example, Mchombo (1978, pp. 290–292) points out that a relative clause modifying the object of the verb, as in (70), simply cannot appear in such expressions:

 (70) nyengá msúngwána améné ánábwérá dzulo → seduce the girl who came yesterday
 \*mnyenga msúngwána améné ánábwérá dzulo 'the-girl-who-came-yesterday seducer'

Note that such expressions remain ungrammatical even when the meaning is purely generic:

<sup>&</sup>lt;sup>43</sup> Carstens (1991, pp. 66–7) suggests that the compounds cannot have semantically specific constituents because they characterize individuals by their professions, habitual activities or attributes. This proposal leaves unexplained why *phrases* that characterize persons by their habitual actions, etc., may have specific constituents, as in *a habitual eater of them* (e.g., *pumpkins*); the typical climber of THIS (is really tough), since, on Carstens' analysis, the constructions are phrases, not morphological words. The proposal also fails to account for the fact that there exist morphological derivatives of specific proper nouns: *McCarthyte*, *Schroederesque*, *New Yorker*, *Tyson-defender*, *Gorbachev look-alike*, *Elvis impersonator*. Liberman and Sproat (1992, pp. 157–8) discuss many other examples.

# (71) nyengá átsíkána ókóndá maphwando → seduce girls who love parties \*mnyenga atsíkána ókóndá maphwando girls-who-love-parties seducer

The infinitives/gerunds, in contrast, allow free modification in all these instances.

Several researchers have assumed, as have Myers (1987), Kinyalolo (1991), and Carstens (1991), that if synthetic compounds contain phrasal strings of words, then they must be generated by syntactic phrase structure principles. Thus, Myers (p.c. June 5, 1989) cites examples from Chishona praise poetry that apparently violate recursivity, such as (72):

(72) mu-gona ku-ronga-nhaú
 1-be able 15-arrange-matter
 one able to settle disputes

Kinyalolo (1991, pp. 222–8) cites Kilega synthetic compounds which include adverbs and adjectival secondary predicates, and Carstens (1991, p. 61) cites similar examples in Kiswahili, such as *mlala uchi* 'one who sleeps naked'. In both cases the postverbal predicate is a VP complement, and not an attributive modifier of nominals.<sup>44</sup> Similarly, Carstens (1991, p. 2) cites instances of modifiers within the compounds, such as *mwandika vitabu vya mapenzi* 'a writer of romantic books', and other examples with conjoined subphrases, such as *mchimba kisima na kaburi* 'a well and grave digger'.

Such cases are found in compounds in English as well. Example (73) shows VP adverbs and predicates, example (74) shows modifiers, and (75) shows conjunctions:

- (73) green-looking, a good-looker, an early riser, stupid-seeming, a go-getter
- (74) American history teacher, raw oyster eater, acoustic guitar player
- (75)a. Charles and Di syndrome, pipe and slipper husband (Lieber 1988, p. 206; 1992, p. 11) peanut butter and jelly sandwich
  - b. Now that my ladder's gone I must lie down where all the ladders start

<sup>&</sup>lt;sup>44</sup> Such examples are ungrammatical in Chicheŵa, which lacks VP complements of this type.

In the foul *rag-and-bone* shop of the heart. (W. B. Yeats, "The Circus Animals' Desertion" [italics added (JB/SAM)])

We have argued, however, that the presence of phrasal strings of words within words is insufficient evidence for the generation of word structure by syntactic principles. Morphological derivatives may be formed from the (possibly innovative) lexicalizing of phrases. The earmarks of this lexicalization are lexical gaps and the existence of lexical integrity properties. Examples like (70)-(71) are evidence of lexical gaps which cannot be explained if these phrases are derived by true phrasal recursivity. We have already seen lexical integrity properties in the failure of extractibility from compounds and their inbound anaphoric islandhood.<sup>45</sup>

Both the conjoinability and gapping tests support the morphological analysis of the synthetic compounds. Concerning gapping, observe that both verbs and nouns can be gapped from syntactic phrases in Chicheŵa, as shown in (76) and (77):

bwăto koma m-lenje (76)M-sodzi a-na-sém-á 1-fisher 1SB-REC PST-carve-IND 15canoe but 1-hunter m-thîko. (a-na-sém-á) (1SB-REC PST-carve-IND) 3-ladle The fisher carved a canoe, but the hunter (carved) a ladle. (77)zi-thúnzí zi-tâtu z-á Chátsalĭra ndí (zí-thúnzí) 8-picture 8-three 8-ASSOC Chatsalira and (8-picture) zí-sanu z-á Joza. 8-five 8-ASSOC J. three pictures of Chatsalira and five (pictures) of Joza

But nouns and verbs cannot be gapped from the nominalized VPs. In (78), for example, the verb *khǎla* 'sit' cannot be gapped from the right conjunct:

(78) m-khala pa chǐtsa ndí ḿ-\* (khala) pa mwǎla 1-sit on treestump and 1- (sit) on rock

<sup>&</sup>lt;sup>45</sup> Although the data given by Kinyalolo (1991) and Carstens (1991) are not completely comparable to ours, it appears that Kilega and Kiswahili synthetic compounds also show lexical integrity properties. Myers (1987) does not investigate lexical integrity properties in Chishona synthetic compounds.

stump-sitter and rock-\*(sitt)er

Again, this evidence supports the morphological analysis of the classifier in (67) over the syntactic analysis in (62).<sup>46</sup>

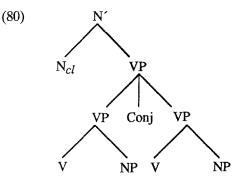
As for conjoinability, we note that in Chicheŵa, conjunction is crosscategorial as it is in English. Not only NPs, but all XP categories can be conjoined:

- (79)a. wá-m-táli kwámbîri ndí wá-m-kâli (AP Conj AP)
   1ASSOC-1-tall very and 1ASSOC-1-fierce
   very tall and fierce
  - b. ndí nkhwângwa kapéná ndí khâsu (PP Conj PP)
     with 9axe or with 5hoe
     with an axe or with a hoe
  - c. A-ku-fúná kutí tí-b-é njínga (CP Conj CP) 1-PRES-want that I SB-steal-SUBJN 9bicycle
    S/he wants that we steal a bicycle
    kapéná kutí tí-gumul-é nyúmba. or that I SB-demolish-SUBJN 9house
    or that we demolish a house.

Given these facts, an ad hoc stipulation would be required to exclude the coordination of VPs.

Now, the syntactic analysis of the synthetic compounds would predict that the inner phrases which are syntactic sisters to the noun class marker should be conjoinable as in (80):

<sup>&</sup>lt;sup>46</sup> Carstens (1991, pp. 66–67) accepts our conclusion that the class marker is a prefix and proposes that the stranded prefix requires raising of the verb to support it. It is stipulated that a null syntactic category cannot host a prefix. This is problematic for the reasons already given in section 5 above.



But it is impossible to conjoin VPs within these nominals, as shown by the ill-formedness of (81b) and (82b):

- (81)a. m- [konda ma-ûngu] ndí m- [pinga njira]
   1 love 6-pumpkin and 1 block 9path
   pumpkin lover and pathblocker
  - b. \*m- [konda ma-ûngu ndí pínga njira] 1 love 6-pumpkin and block 9path
- (82)a. m- [pinda ma-lâta] kapéná m- [pala ma-tǎbwa]
   1 bend 6-corrugated iron or 1 scrape 6-timber/plank
   metalworker or carpenter
  - b. \*m- [pinda ma-lâta kapéná pala ma-tǎbwa] 1 bend 6-corrugated iron or scrape 6-timber

Thus, by the conjoinability test the nominalizations behave like morphological words and not syntactic phrases.<sup>47</sup> Even though conjunctions can occasionally appear within lexicalized phrases (as in English *rag-and-bone shop of the heart* or Kiswahili *mchimba kisima na kaburi* 'well and grave

(i) \*m-[phwanya ndí kázinga] ma-zira 1-crack and fry 6-egg cracker and fryer of eggs

<sup>&</sup>lt;sup>47</sup> Carstens (1991, pp. 69–70) proposes that the ungrammaticality of these examples follows from the fact that the prefix requires a morphological host, thus accepting our conclusion that it is a morphological prefix, together with the assumption that the verb stem must raise to support the prefix, but that it cannot do so out of a conjoined VP. This proposal fails to explain why a conjoined V cannot raise to support the prefix in (i):

Both mphwanya mazira 'egg cracker' and mkazinga mazira 'egg fryer' are possible compounds.

digger'), it remains true that the class markers cannot be stranded by conjoining the morphological constituents to which they are prefixed.

Our results thus far support Myers' (1987) syntactic analysis of the infinitives/gerunds, as in (58), and Mchombo's (1978) and Sproat's (1985) morphological analysis of the synthetic compounds, as in (67). However, other predictions of Myers' syntactic analysis of infinitives/gerunds are *not* borne out. Unlike the word-external classifiers that we have already seen, class 15 ku- cannot be stranded by conjunction in (83)–(84) or gapping in (85):

- (83) \*ku- [píndá njǐnga ndí/kapéná pala ma-tǎbwa]
   15- bend 10bicycle and/or scrape 6-timber/plank
   to repair bicycles and/or do carpentry
- (84) \*ku-[phwânya ndí kázinga] ma-zira
   15-crack and fry 6-egg
   to crack and fry eggs
- (85) ku-píndá njǐnga ndí/kapéná ku-\* (píndá) ndowa
   15-bend 10bicycle and/or 15- (bend) 10pail/bucket
   to repair bicycles and/or metal pails

Stranding is predicted by the syntactic analysis, according to which kuclassifier takes a VP sister. In (83), two such VPs are conjoined under the same classifier.<sup>48</sup> In (84), the head V within the VP is a syntactic conjunct.

<sup>&</sup>lt;sup>48</sup> Note that the class 15 prefix, unlike the prefixes of synthetic compounds, does not carry number information (although like these prefixes, it does induce class 15 concord on nominal modifiers and class 15 subject and object agreement on the verb). Consequently, Kinyalolo's (1991, p. 231) proposal that the class prefix is an inflectional spell-out of the number feature on the head N, which therefore spreads over conjuncts, is not applicable in this case. Still other assumptions must be added to account for these facts on the head-movement theory.

Carstens (1991, p. 209) hypothesizes that class 15 concord in Kiswahili "POSS-ing" kuphrases is actually class 17 "expletive" concord. This proposal for explaining class 15 concord cannot work in general. Chicheŵa lacks expletive uses of the locative concord (Bresnan and Kanerva 1989). Moreover, in Chicheŵa, the ku- of class 15 and the ku of class 17 are distinguished morphologically: only the former undergoes contraction to o following the associative 4-. Many other Bantu languages distinguish classes 15 and 17. Kichaga, for example, has preserved the locative concords ha- and ku- for classes 16 and 17 but employs class 5 *i*- for infinitives/gerunds and their concords (Lioba Moshi, personal communication, June 17, 1992). In Sesotho the locative class prefixes have largely disappeared; in particular, there are no locative concords for adjectives or verbal objects. But the class 15 concords for adjectival modifiers and verbal object concords of infinitives/gerunds are retained ('Malillo Machobane, personal communication, June 17, 1992). Gitonga similarly preserves infinitive/gerund (class 15) but not locative concords (Lanham 1955). Consequently, the strategy of trying to explain class 15 concord as class 17 concord cannot work in general.

In (85), two NPs, each headed by ku-, are conjoined, and the repeated verb is gapped from within the phrasal sister of the second. Even if we assume that verbal categories cannot be conjoined in Chicheŵa, an assumption without independent motivation, the failure of gapping within (85) remains unexplained.

In failing to strand under gapping and conjunction, ku- behaves like a verbal prefix rather than a nominal head. This result is not surprising when we consider that, like noun stems, bare verb stems in Bantu are morphologically bound forms. Apart from their use as the imperative form of the verb, bare verb stems can never appear without prefixes. They cannot be predicated of anything, conjoined with anything or embedded as the complements of any lexical items. This property of the infinitive/gerund is not captured by Myers' syntactic analysis (58). In fact, there are further syntactic properties of infinitives/gerunds in Chicheŵa that are inconsistent with the analysis of ku- as a nominal head, as in Myers' analysis.

Among the verbs which take infinitive/gerund phrase complements are those which are exclusively transitive, like *zonda* 'hate' in (65) and (86a). In (86b), the verb takes a simple NP complement, while (86c) shows that this verb can take an object marker corresponding to an infinitive. Examples (86d,e) show that the infinitival complement of this verb can have nominal modifiers.

- (86)a. A-na-zónd-á ku-bwêra ndí kú-dyá
  2SB-REC PST-hate-IND 15-come and 15-eat
  nkhúkú y-åthu.
  9chicken 9-our
  S/he hated to come and eat our chicken.
  - b. A-na-zónd-á chi-yâni?
     1-REC PST-hate-IND 7-what
     What did he hate?
  - c. A-na-kú-zónd-a.
    2SB-REC PST-15OB-hate-IND
    S/he hated it (e.g. coming and eating our chicken).
  - d. A-ku-zónd-á ku-píndá njingá uku.
     2SB-PROG-hate-IND 15-bend 10bicycle 15this
     S/he hates this bending bicycles.

e. A-ku-zónd-á ku-píndá njingá kw-áko.
 2SB-PROG-hate-IND 15-bend 10bicycle 15-your
 S/he hates your bending bicycles.

All of these properties are consistent with the nominal character of infinitives/gerunds.

However, there is also a class of intransitive auxiliary-like verbs which take infinitival complements, and these contrast in all respects to transitive verbs like *zonda* 'hate' in (86). An example is *yeněra* 'have to' in (87). Compare (87a-e) with (86a-e):<sup>49</sup>

- (87)a. A-na-yénér-á kú-bwêra ndí kú-dyá
  2SB-REC PST-must-IND 15-come and 15-eat
  nkhúkú y-ăthu.
  9chicken 9-our
  S/he had to come and eat at our chicken.
  - b. \*A-na-yénér-á chí-yâni?
    1-REC PST-must-IND 7-what
    Lit.: \*What had he to?
  - c. \*A-na-kú-yéněr-a.
     2SB-REC PST-15OB-must-IND
     Lit.: S/he had to it.
  - d. \*A-ku-yénér-á kú-píndá njingá uku.
     2SB-PROG-must-IND 15-bend 10bicycle 15this.
     Lit.: S/he has to this repairing bicycles.
  - e. \*A-ku-yénér-á kú-píndá njingá kw-áko.
     2SB-PROG-must-IND 15-bend 10bicycle 15-your
     Lit.: S/he has to your repairing bicycles.

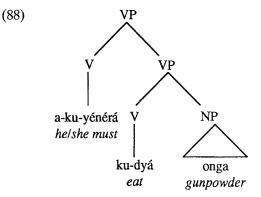
These verbs can take only infinitive phrases, as in (87a), and cannot take NP complements at all, even those that refer to an infinitive phrase, as

<sup>&</sup>lt;sup>49</sup> Tabor (1992) discusses further contrasts between the two classes, including coalescence of a + ku to o in the Kasungu/Nkhotakota dialect of Chicheŵa. The verb *konda* 'like' in (60) is ambiguous.

in \*Akuyénérá uku 'S/he has to do this (something of class 15, e.g., repairing bicycles)'.<sup>50</sup>

Note that the complements of both the intransitive verbs like *yeněra* 'have to' and the transitive verbs like *zonda* 'hate' behave identically with respect to all of the lexical integrity tests. In particular, the ungrammaticality of stranding ku- through conjunction or gapping as in (83) and (85) is preserved when the infinitive is embedded under either verb. Also, the infinitival complements to both verbs allow extraction, as in (65).

We see then that, contrary to Myers' syntactic analysis, the ku + infinitive phrase can appear in non-nominal positions having the essential structural properties of (88).<sup>51</sup> What is essential in (88) is that the verb is prefixed by the infinitive marker, and the entire verbal phrase is nonnominal. (It does not matter for our purposes here exactly which nonnominal category it is, VP, S, or IP, that dominates the infinitive phrase.)



Thus, the nominal/verbal duality of the infinitive/gerund construction cannot be traced to the syntactic status of class 15 ku- as a nominal head taking a VP complement, contrary to Myers' analysis. In all cases, class 15 ku- behaves like a verbal prefix. How then can the categorial duality be explained?

One approach is to give up a uniform analysis of ku- and not attempt to relate the (verbal) infinitive marker in (87) and (88) to the (nominal) class marker in (86). The categorial duality of infinitives/gerunds would be a reflection of an ambiguity of ku-. This approach leaves most of the

 $<sup>^{50}</sup>$  In Chishona there is a similar class of verbs, but the infinitive *ku*- marker has coalesced with the final vowel of the auxiliary, leading Myers (1987, pp. 89–95) to propose that reanalysis of his hypothesized infinitive structure has taken place so that the auxiliaries take bare VP complements. The existence of conjunctions, as in (87a), shows that no such reanalysis has applied in Chicheŵa.

<sup>&</sup>lt;sup>51</sup> Note that it follows from this result together with (87a) that VPs (or extended nonnominal projections of verbal phrases) can be conjoined in Chicheŵa.

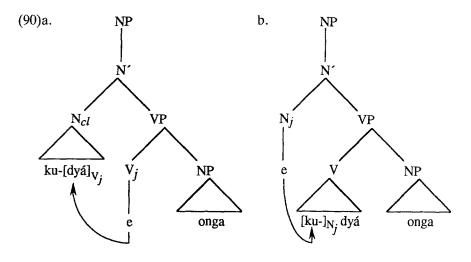
problem unsolved, however. It fails to explain the shared properties of the constructions in complement selection, extraction and gender class. Moreover, there is evidence that even the 'verbal' ku- in (88) carries the class 15 gender information. In the following example, the class 15 object marker of the verb *hate* agrees with the infinitival complement of the verb *must*. In Chicheŵa object markers show anaphoric agreement in noun class with their antecedents (Bresnan and Mchombo 1987).

(89) A-ku-yénérá kú-dy-á onga, koma 1SM-PROG-must-FV 15-eat-FV 14gunpowder but á-ma-ku-zônd-a.
1SM-PRS HAB-15OM-hate-FV S/he has to eat gunpowder but s/he hates it (eating gunpowder).

Note that the object marker ku- cannot refer to gunpowder, which is of class 14; to make that reference, the class 14 object marker u- must be used. Nor can ku- refer to a finite 'that' (kuti) complement, which lacks the class 15 feature (Bresnan and Kanerva 1989, Bresnan 1994b).

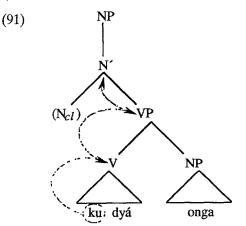
Another approach is to assume that the categorial duality of gerund phrases is captured by means of syntactic movement. In the syntax a VP headed by the verb stem is embedded in an NP projected or "converted" (Abney 1987, pp. 223–4) by the verbal affix occupying a syntactically independent position in phrase structure. The stem and its affix are joined syntactically through head-raising in (90a), affix-lowering in (90b), or affix-lowering in P(honetic)F(orm).<sup>52</sup>

 $<sup>^{52}</sup>$  See Abney (1987, ch. 3) for a review of such analyses and their various theory-internal implications.



However, neither of these analyses accounts for the nonnominal character of the infinitives of the type shown in (88) because they both assume that the *ku*- prefix originates as a syntactic NP constituent. Further, neither analysis explains the lexical integrity properties of the derived syntactic unit consisting of the verb stem and affix. We have seen, for example, that  $X^0$  categories can be conjoined in Chicheŵa, including the locative classifiers (47a), and that syntactically independent classifiers can be stranded by conjunction and ellipsis. Prefixes, however, cannot be, and nothing in their syntactic analysis as  $X^0$  categories in head movement or affix lowering configurations explains this.

An alternative approach assumes that the verb stem is affixed in the morphology in accordance with the lexical integrity principle, but that the affix carries nominal features which allow the syntactic projections of the infinitive/gerund to be embedded in nominal phrases, as illustrated in (91):



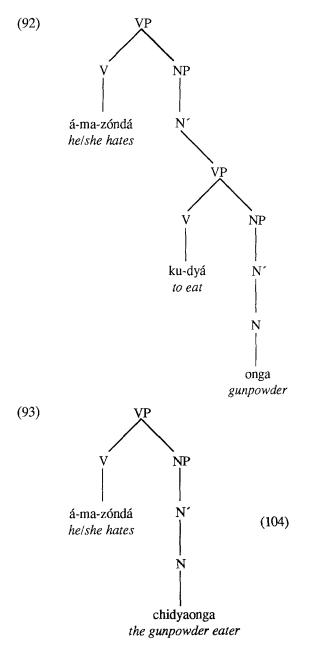
Note how the dependency between the presence of the ku- prefix on the verb and the infinitival VP's being embeddable in a nominal phrase is captured. All NPs in Chicheŵa must have a noun class. Of all the noun class markers, only ku- derives verbals, contributing features such as noun class, which are essential to nominal phrases. Hence, a VP headed by ku-V is the only type of VP that can head a nominal construction.

This analysis violates categorial endocentricity because the gerund phrase is treated as an exocentric phrase in phrase structure, a headless NP dominating a VP. A technical solution to this problem would be to postulate an 'abstract element' KU to occupy the head of N' position whose features must match the features of the morphologically attached ku-morpheme.<sup>53</sup> But because no lexical elements ever occupy this syntactic position, unlike other positions where X' heads are motivated, the analysis preserves endocentricity in name more than in substance.

A more interesting proposal exploits the use of parallel levels of structure at which 'head' is defined in our framework. From the point of view of the surface categorial structure, the NP in (91) is exocentric because its categorial head is not found within it. But from the point of view of the parallel level of f-structure, the head of the NP is the gerund itself, the lexical head of VP. This seems to be what permits the violation of categorial exocentricity (Bresnan 1982; Kroeger 1993).

Summing up this section, the infinitive/gerund is confirmed by our tests to be a syntactic phrase, except for its classifier ku-, which is a verbal prefix as in (92). In contrast, synthetic compounds, though they can contain phrasal strings, have the earmarks of morphological entities. They are formed by prefixation of the noun class marker to (innovatively) compounded or lexicalized phrases, as in (93).

 $<sup>^{53}</sup>$  Just such an analysis of the English gerund phrase is attributed by Abney (1987, pp. 241– 4) to Chomsky in personal communication. It could be modelled in LFG by imposing a feature constraint as an alternative to the expansion of N before VP in (91).



These contrasts are inexplicable under the syntactic analyses. According to these analyses, both constructions contain syntactic projections of verbs. After verb movement, the residue of complements in the VP of both phrases should show syntactic properties. Nothing in the syntactic theory of word formation explains why only one of the two identical phrasal configurations has the expected syntactic properties. Nor do the syntactic analyses explain the lexical integrity properties of the prefixed verbal forms themselves, as we saw above.

## 7. CLASS-MARKED ADJECTIVES: DEADJECTIVAL NOUNS

There is one final use of the noun class marker that is highly revealing for our investigation. From Bleek (1862) onward, the derivational nature of the Bantu noun class markers has been repeatedly noted, and it is a primary motivation for the traditional Bantuist analysis of the class markers as prefixes.<sup>54</sup> For example, certain noun class markers are used to derive nouns from adjective stems (Mufwene 1980). In Chicheŵa, too, the class 14 prefix is used to derive nouns from a distinct lexical category of adjectives and numerals. The derived nouns themselves may undergo further derivational morphological processes. Hence, the class marker prefix must also be a morphological and not a syntactic element. This conclusion follows from the generalization that derivational morphological processes operate on lexical units such as stems and affixes, as implied by the lexical integrity principle. There is no account for these results on the syntactic analyses, short of importing derivational morphology into the syntax.

As in other Bantu languages, adjectives are a closed class in Chicheŵa. Adjective stems include the following shown in the table in (94).

(94	

Adjective	e Stems:
-tâli	long short
-fûpi -ng'ôno	snort small
-kûlu -wîsi	big
-wisi -kâli	raw, unripe, immature fierce, sharp

Adjectives have several properties that clearly distinguish them from nouns. First, they differ in morphemic structure. As (95) shows, though an adjective stem bears the same noun class prefix as a noun, it requires an additional prefix in Chicheŵa, the 'associative prefix' (Orr and Scotton 1980), which doubles the gender class information:

 $<sup>^{54}</sup>$  See the references cited in n. 7.

(95)

Morphemic Structure:		
Noun:	Adjective:	
[prefix <sub>1</sub> + stem] m – kâzi woman (class 1)	$[prefix_{1} + stem]$ *m - wîsi <i>immature</i> (class 1) $[prefix_{assoc} + [prefix_{1} + stem]]$ w-á - m - wîsi <i>immature</i> (class 1)	

Second, adjective and noun stems differ in the range of classes they select. Most noun stems select only a few classes, while adjectives occur in all classes, as shown in (96):

(96)

Noun:		Adjective:	
m-kâzi wa-kâzi *li-kâzi *zi-kâzi	woman (cl. 1) women (cl. 2) ? (cl. 5) ? (cl. 8/10)	wá-m-wîsi á-á-wîsi la-lí-wîsi zá-zí-wîsi	immature (cl. 1) immature (cl. 2) unripe (cl. 5) raw (cl. 8/10)
	•		•

Third, adjectives reduplicate differently from nouns. Adjectives can reduplicate by copying the entire adjective minus the outer associative prefix. This reduplicated form has the meaning of distributing the adjectival property A over a group or mass. An adjective A reduplicated in this way could be translated as 'all being A'. This is illustrated in (97).

Adjective Reduplication:		
wá-m-wîsi	1ASC-1-short	wá-mwísĭ-mwĭsi
á-á-kûlu	2ASC-2-big	á-ákúlú-akûlu
yá-í-ng'ôno	4ASC-4-small	yá-íng'onó-ing'ôno
lá-lí-tâli	5ASC-5-long	lá-lítálí-litâli
chá-chí-fûpi	7ASC-7-short	chá-chífúpí-chifûpi
zá-zí-kâli	8ASC-8-sharp	zá-zíkálí-zikâli

This adjectival reduplication pattern is ungrammatical with nouns: \*lamulo-lamulo 'law'-REDUP, \*madengu-madengu basket-REDUP, etc.<sup>55</sup>

Fourth, adjectives differ from nouns in disallowing preprefixes (Doke 1935, p. 43).<sup>56</sup> This is shown in the table in (98), in contrast to (29).

(98)

No Preprefixation of Adjective Stems:
[prefix <sub>2</sub> + stem] ka – wîsi immature (class 12, diminutive)
[prefix <sub>2</sub> + [prefix <sub>1</sub> + stem]] *ka - m - wîsi diminutive (class 12) immature one (class 1)

Fifth, bare adjective stems can take locative class prefixes (classes 16, 17, 18), while bare noun stems cannot, as pointed out by Doke (1935, p. 43) for Bantu in general.<sup>57</sup> See the table in (99).

<sup>&</sup>lt;sup>55</sup> Nouns only reduplicate by copying the final disyllabic foot of the noun and adding a high tone (Kanerva 1990, pp. 51–53). The meaning is intensive rather than distributive. Thus a reduplicated noun N could be translated 'a real N'. Adjectives can also follow the noun reduplication pattern, but only with the intensive meaning 'really A' and not the distributive meaning 'all being A'.

<sup>&</sup>lt;sup>56</sup> An associative prefix must be added to the unstarred form in this table to complete the derivation of the adjective, which in this case will be  $k\dot{a}$ - $k\dot{a}$ - $w\hat{s}\hat{s}i$ .

<sup>&</sup>lt;sup>57</sup> As in the previous example, the outer associative prefix is omitted from th adjective in this table to highlight the contrast to the noun.

1	n	$\gamma$
ł	72	")

Position of Locatives in Morphemic Structure:	
Adjectives	Noun:
[loc + stem] pa – wîsi immature (class 16, locative)	<pre>[loc + stem] *pa - kâzi woman (class 16, locative) [loc [<sub>N</sub>prefix<sub>1</sub> + stem]] pa m - kâzi on (class 16) the woman (class 1)</pre>

There are a few lexical exceptions to this generalization: noun stems that directly take locative prefixes include *m-kǎti* 'inside', *pa-kǎti* 'between', *pǎ-nja* 'outside (of something specific)', *kǎ-nja* 'outside'. It is clear that *-kǎti* and *-nja* are noun stems and not nouns with a phonologically null prefix because (1) they are bound forms, and (2) they do not allow alternative concord. For example, in *pa-njá pó-yêra* 'a clean yard', the class 16 locative *pǎ-nja* 'outside, yard' can be modified by the class 16 attributive *pó-yêra* 'clean', but not by a class 9 modifier: \**pa-njá yó-yêra*.

Altogether, these five properties sharply distinguish adjective stems from noun stems.<sup>58</sup> Thus, adjectives do belong to a different lexical category from nouns. However, nouns can be derived from adjectives by prefixing the class 14 noun class marker to the bare adjective stem.<sup>59</sup>

<sup>&</sup>lt;sup>58</sup> Adjective stems can also be distinguished from verb stems. First, they fail to take the typical inflectional morphology of verbs. Second, they do not allow verb derivation suffixes such as the causative, passive, applicative, reversive, stative, reciprocal or intensive. Third, verb reduplication differs from reduplication of adjectives and nouns in that no prefixes can be reduplicated with the verb stem (Mtenje 1988; Kanerva 1990).

<sup>&</sup>lt;sup>59</sup> The same prefix derives nouns denoting states or qualities from noun stems (e.g., *u-kâzi* 'womanhood'). A few deadjectival nouns have been also derived by other noun class prefixes than class 14: *li-kûlu* 'headquarters, the capital (class 5)', *m-ng'ôno* 'younger sibling of the same sex (class 1)', *m-kûlu* 'older sibling of the same sex (class 1)'.

Nominalized Adjectives		
u-tâli	14-long	length
u-fûpi	14-short	shortness
u-ng'ôno	14-small	smallness
u-kûlu	14-big	bigness, size
u-wîsi	14-raw, unripe, immature	rawness, ripeness, immaturity
u-kâli	14-fierce, sharp	fierceness, sharpness

The deadjectival nouns in (100) show all of the expected syntactic characteristics of nouns. For example, they induce class 14 concord in nominal modifiers as well as class 14 subject-verb agreement, as illustrated in (101a,b):

- (101)a. U-kúlú w-á nyúmbá iyi
  14-big 14-ASC 9house 9this
  u-ku-tí-ópsy-a.
  14SB-PROG-I PL O-frighten-IND
  The size of this house frightens us.
  - b. U-wísí uwu u-ku-tsímíkiz-a kutí 14-unripe 14this 14SB-PROG-indicate-IND that chí-pátso=chi chi-sa-dy-êdw-e. 7-fruit=7this 7SB-NEG-eat-PASS-SUBJUN

This unripeness indicates that the fruit should not be eaten.

These deadjectival nouns show the expected morphological properties of nouns as well. They do not require the associative marker, as shown in (101). They take preprefixes, such as the augmentative *chi-: chi-ukûlu* 'great size'. They allow locative markers outside their class 14 prefix: *pa ukûlu* 'on greatness'. And they allow intensive but not distributive reduplication:<sup>60</sup>

<sup>&</sup>lt;sup>60</sup> The tonal patterns of these forms match those for nonderived nouns such as *mwamú-ná-mŭna* 'man (intensive)'. See Kanerva (1990, ex. (120), p. 195).

1	1	nn
1		UZ.E
۰.		0

Nominalized Adjective Reduplications			
u-tâli u-fûpi u-ng'ôno u-kûlu u-wîsi	length shortness smallness bigness, size rawness, unripeness, immaturity	utálí-tăli ufúpí-fŭpi ung'ónó-ng'ŏno ukúlú-kŭlu uwísí-wĭsi	
u-kâli	fierceness, sharpness	ukálí-kăli	

Distributive reduplication, which is characteristic only of adjectives, is ungrammatical: ufupfufupi 'shortness (distributed)', ukuluukulu 'bigness (distributed)'.<sup>61</sup>

As we would expect of a morphological process of word formation, the nominalizations show lexical integrity properties (insofar as our tests are applicable). The classifier plus stem behaves like a morphological word with respect to tests such as conjoinability and gapping: \*u-tâli ndi -fûpi 'length and shortness', \*u-kûlu kapéná ng'ôno 'bigness or smallness'.

Finally, these morphologically derived nouns undergo other processes of derivational morphology. We have just seen above that they undergo intensive noun reduplication. Among the processes of noun derivation found in the Bantu languages are those that derive names from common nouns by prefixation. Proper names in Isizulu, for example, may be formed from nouns by eliding the initial vowel and prefixing the Class 1a prefix *u*- (Doke 1988, p. 69): *uSikhotha* 'Sikota (class 1a, name)' from *isikhotha* 'grass (class 7)'; *uNtaba* 'Ntaba (Class 1a, name)' from *inthaba* 'mountain (class 9)'; etc. Class 14 deadjactival nouns can also undergo this process.<sup>62</sup>

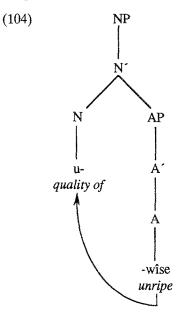
- (103)a. Ubu-bi bu-ya-ngi-khohlisa.
   14-evil 14SM-T/A-me-deceive
   Evil deceives me
  - b. U-bu-bi u-ya-dansa.
    1a-14-evil 1aSM-T/A-dance
    Ububi [name] is dancing.

<sup>&</sup>lt;sup>61</sup> Adjective reduplication with class 14 concord produces similar forms, but with a purely adjectival meaning: *u-dzú wá-ú-fúpí-u-fûpí* 'grass all being short'.

<sup>&</sup>lt;sup>62</sup> These are the judgments of Vuyo Booi, a bilingual speaker of Xhosa and Isizulu from Ladyfrere in the Transkei, South Africa.

In Sesotho, masculine or feminine proper nouns can be formed by prefixing ra- or 'ma-, respectively, to nouns (Doke and Mofokeng 1985, p. 93): Rabōròkò 'Father-of-sleep (class 1a, name)' from bōròkò 'sleep (class 14)', 'Malēkata 'Mother-of-closure (class 1a, name)' from lēkata 'closure (class 5)'. Note that  $b\bar{o}r\partial k\dot{o}$  'sleep' consists of the class 14 marker  $b\bar{o}$ -prefixed to the stem  $r\partial k\dot{o}$  (cf. mar $\partial k\dot{o}$  'evil dreams (class 6)' (Doke and Mofokeng 1985, p. 80)). In Chicheŵa, too, there is a prefix Ná- which forms proper names from common nouns. The name Námágetsi is derived from ma-getsi 'electric lights (class 6)', the name Náúlěsi from u-lěsi 'laziness (class 14)', and Náúwîsi from u-wîsi 'unripeness (class 14)'. The last example is a deadjectival noun converted into a proper noun.<sup>63</sup>

What would a head-movement analysis of these facts look like? The most straightforward account would have class 14 u- analyzed as a noun that takes an AP complement, whose head could then incorporate into the higher noun:



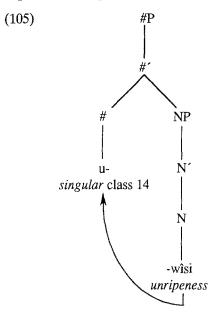
Such an account is implausible, though, because nouns do not select AP complements elsewhere in Chicheŵa, and there are no AP modifiers that can be stranded by the A-movement.<sup>64</sup>

Carstens (1991, pp. 40-41, 78) proposes a variation of this syntactic

<sup>&</sup>lt;sup>63</sup> This derivational prefix, as we would expect, does not allow inner concord with its noun and cannot be stranded by conjunction and ellipsis.

<sup>&</sup>lt;sup>64</sup> This analysis and the arguments against it were suggested by a reviewer.

analysis. On her analysis, u- is the class 14 allomorph of the singular feature that heads a number phrase functional projection in the syntax. The adjective stem  $-w\hat{s}i$  'unripe' is nominalized by lexical affixation of a null morpheme  $\emptyset$  '-ness' of class 14. This nominalized stem projects the inner NP within the '#P' in (105). Syntactic head-movement then incorporates this nominal stem with its syntactic number feature, yielding uwisi 'unripeness' as a spell-out:



Carstens' proposal factors apart the derivational and inflectional aspects of this class marker by treating gender as a lexical feature of stems and number as a syntactic feature appearing in the head position of the number phrase (#P), into which the stem must move to receive its prefix *u*-. As we have already noted (see n. 35), this account requires the duplication of null morphemes in the syntax and morphology.<sup>65</sup> It is clear that even

 $<sup>^{65}</sup>$  In Chicheŵa, the class 14 prefix can also be used as a preprefix, leading to the same problems we previously noted with class 12 ka- in n. 35. For example, we have *u-munthu* 'humanity (class 14)' from *mu-nthu* 'person (class 1)', *u-mwǎna* 'childishness (class 14)' from *mw-ǎna* 'child (class 1)', *u-mbǔli* 'ignorance' from *mbǔli* 'an ignorant person (class 9)'. To derive these forms on the lexical account, one need only say that class 14 *u*- can be a stemlevel or word-level prefix for deriving abstract nouns. On Carstens' head-movement account, however, the meaning of abstraction comes not from the prefix, which is just a morphological realization of syntactic number, but from a null lexical affix which derives class 14 noun stems from bare adjective or noun stems. However, this hypothesized lexical affix cannot explain the meaning of abstraction in the preprefixed class 14 nouns, which can only be derived from syntactic phrase configurations on this account. For such cases, Carstens

on this account, the derivational and inflectional properties of the class marker cannot be neatly factored apart into lexical and syntactic sources. The class marker is a marker of syntactic number, originating in a phrase structure position (#P), yet it appears inside the noun-derivation prefix (*Na*- in Chicheŵa). Thus, the head-movement approach cannot be confined to inflectional morphology but requires syntactic analyses of purely derivational morphology as well.

### 8. CONCLUSION

Let us return to the question that motivated our inquiry at the beginning of this study: How can we tell whether a sequence of morphemes is a word? In the domain of Bantu linguistics, this is a pressing question. As we have seen, the Bantu noun class markers straddle the borderlines between syntax and morphology and between inflection and derivation. They have invited conflicting analyses in both the earliest missionary grammars and the most recent theoretical works in generative grammar. For these reasons, they are a particularly rich and fruitful domain for the investigation of lexical integrity.

In the architecture of syntactic theories of word formation, information about syntactic functions and relations is expressible only in the vocabulary of syntactic phrase structure. In that framework, it is natural that the inflectionally rich word structure we find in Bantu should be transformationally derived from syntactic phrases, in violation of the lexical integrity principle. However, the evidence we have found indicates that Bantu words are indeed built out of different structural elements and by different principles of composition than syntactic phrases, in accordance with the lexical integrity principle, even though these words and phrases may carry the same information about syntactic functions and relations. The LFG architecture which provides our general conceptual framework makes sense of these results by factoring apart the representation of syntactic functions and relations (the 'inner' structure of language) from the representation of structural formation (the 'outer', or surface, structure). These are parallel information structures, which are related not by transformational derivation, but by functional correspondences.<sup>66</sup>

Because in this framework the correspondence between structural form

hypothesizes another null morpheme – this one occupying a syntactic position outside of the #P from which the inner prefixed noun is derived by head movement, but having the same meaning and gender class as the lexical affix.

<sup>&</sup>lt;sup>66</sup> See Bresnan and Mchombo (1987) and Bresnan (1994a) for explicit analyses of Bantu in this framework.

and syntactic function is in general imperfect, changes in form can occur partly independently of changes in function. And this, we hypothesize, is just what we are seeing in Bantu. The Bantu noun class markers have been in the midst of a historical change from syntactic constituents to bound morphemes. When a given class marker has crossed the boundary between phrase and word, many of its grammatical properties are systematically changed, as shown by tests derived from the lexical integrity principle. Despite these systematic differences between the word-internal and phrasal markers, their functions in grammatical system of concord are identical. The differences depend only on whether the class markers are syntactic or morphological constituents. In other words, the minimal difference between the two types of class markers is just the syntax-morphology boundary itself. This provides striking support for the lexical integrity principle.

It is natural to ask<sup>67</sup> whether the lexical integrity principle cannot be built into the syntactic head movement theories by means of morphological well-formedness constraints on the  $X^0$  category. The lexical integrity principle, under this interpretation, would be a kind of output filter applying to the structures built by head movement. It is important to understand why this approach to lexical integrity fails to account for our results. The reason is simply that there is no natural structural characterization of the word in terms of syntactic categories. Among  $X^0$  categories, some are formed by morphological processes, but others are syntactically basegenerated, including conjoined  $X^0$ s, periphrastic verbs (Poser 1991) and other structures (Sells and Cho 1991). These syntactic  $X^0$  structures do not show the strong lexical integrity properties we have identified in morphological words. Thus, the label  $X^0$  does not suffice to identify the morphological word in syntactic terms.

Moreover, several of our most telling arguments for lexical integrity depend not on the category built by syntactic head movement, but on the hypothesized phrasal domains of such movement – which would *not* be dominated by  $X^0$  at any point. One argument is the absence of inner agreement modifiers for preprefixed nouns (discussed in sections 3–5). Such modifiers are dominated by phrasal nodes, and their absence would not follow from well-formedness constraints on  $X^0$ . Another argument is the absence of extraction sites and anaphoric pronouns from the synthetic compounds, in contrast to the infinitives/gerunds (see section 6). These properties appear in domains which are dominated by verbal projections and not by  $X^0$ , according to the syntactic analyses. Still another argument

<sup>&</sup>lt;sup>67</sup> – as did a reviewer.

is the use of the class markers in the derivation of deadjectival nouns in section 7, which may undergo further processes of derivational morphology. Here the problem is to explain how functional projections can be embedded inside derivational morphemes without violating the proposed morphological constraints on the  $X^0$  category. The alternative of giving a syntactic phrasal source to purely derivational morphology only worsens the problem of recapturing the effects of the lexical integrity principle syntactically.

While Bantu provides a particularly rich empirical domain for studying morphology-syntax interactions, its morphosyntactic properties, stacked inflectional morphemes, mixed derivational and inflectional uses of the same morpheme, are not atypical, as a look at the descriptive and typological literature will show. We conclude that the recent trend of hypothesizing syntactic movement approaches to inflectional morphology is wrong for the description of Bantu and as a general theory. However, the rejection of syntactic movement approaches to morphology does *not* require the abandonment of extended X' theories of functional projections as explanations for word order variations across languages. Recent work shows that by means of unification, a constraint-based theory of grammar such as LFG can fully model the *syntactic* effects of verb movement to functional projections while maintaining the strong morphosyntactic constraints of the lexical integrity principle (Netter and Kärcher 1986; Na+ter 1988; Kroeger 1993; King 1992).

#### References

- Abney, Steven: 1987, The English Phrase in Its Sentential Aspect, Ph.D. dissertation, MIT. Ackerman, Farrell: 1987, Miscreant Morphemes: Phrasal Predicates in Ugric, Ph.D. dissertation, University of California, Berkeley.
- Ackerman, Farrell and Gert Webelhuth: 1992, 'Topicalization and German Complex Predicates', unpublished manuscript, UCSD.
- Allan, Keith: 1977, 'Classifiers', Language 5, 285-311.
- Alsina, Alex: 1993a, Predicate Composition: A Theory of Syntactic Function Alternations, Ph.D. dissertation, Stanford University.
- Alsina, Alex: 1993b, 'A Theory of Complex Predicates: Evidence from Causatives in Bantu and Romance', paper presented at the Workshop on Complex Predicates, CSLI, Stanford University, May 21-3, 1993.
- Alsina, Alex and Sam A. Mchombo: 1990, 'The Syntax of Applicatives in Chicheŵa: Problems for a Theta Theoretic Asymmetry', Natural Language & Linguistic Theory 8, 493-506.
- Alsina, Alex and Sam A. Mchombo: 1993, 'Object Asymmetries in the Chichewa Applicative Construction', in Sam Mchombo (ed.), *Theoretical Aspects of Bantu Grammar*, CSLI, Stanford University, pp. 17–45.

Anderson, Stephen R.: 1984, 'Kwakwala Syntax and the Government-Binding Theory', in

Eung-Do Cook and Donna B. Gerdts (eds.), Syntax and Semantics 16: The Syntax of Native American Languages, Academic Press, New York, pp. 21-75.

- Andrews, Avery D: 1990, 'Unification and Morphological Blocking', Natural Language & Linguistic Theory 8, 507–557.
- Baker, Mark: 1988a, Incorporation: A Theory of Grammatical Function Changing, University of Chicago Press, Chicago.
- Baker, Mark: 1988b, 'Theta Theory and the Syntax of Applicatives in Chicheŵa', Natural Language & Linguistic Theory 6, 353-89.
- Baker, Mark: 1992, 'Thematic Conditions on Syntactic Structures: Evidence from Locative Applicatives', in I. M. Roca (ed.), *Thematic Structure: Its Role in Grammar*, Foris Publications, Dordrecht, pp. 23-46.
- Bleek, W. H. I.: 1862, A Comparative Grammar of South African Languages, Trubner and Co., London.
- Booij, Geert E.: 1985, 'Coordination Reduction in Complex Words: A Case for Prosodic Phonology', in H. van der Hulst and N. Smith (eds.), Advances in Nonlinear Phonology, Foris Publications, Dordrecht, pp. 143–160.
- Booij, Geert E. and Jerzy Rubach: 1984, 'Morphological and Prosodic Domains in Lexical Phonology', *Phonology Yearbook* 1, 1–27.
- Booij, Geert E. and Jerzy Rubach: 1987, 'Postcyclic versus Postlexical Rules in Lexical Phonology', *Linguistic Inquiry* 18, 1-44.
- Botha, Rudolf P.: 1980, Word-based Morphology and Synthetic Compounding, Stellenboosch Papers in Linguistics 5, Department of General Linguistics, University of Stellenbosch, Stellenbosch 7600, South Africa.
- Botha, Rudolf P.: 1981, 'A Base Rule Theory of Afrikaans Synthetic Compounding', in M. H. Moorgat, H. van der Hulst and T. Hoekstra (eds.), *The Scope of Lexical Rules*, Foris Publications, Dordrecht, pp. 1–77.
- Botha, Rudolf P.: 1984, Morphological Mechanisms: Lexicalist Analysis of Synthetic Compounding, Pergamon Press, Oxford.
- Bresnan, Joan: 1978, 'A Realistic Transformational Grammar', in M. Halle, J. Bresnan and G. A. Miller (eds.), *Linguistic Theory and Psychological Reality*, MIT Press, Cambridge, pp. 1–59.
- Bresnan, Joan: 1982, 'On Control and Complementation', Linguistic Inquiry 13: 343–434; also in Joan Bresnan (ed.), The Mental Representation of Grammatical Relations, MIT Press, Cambridge, pp. 282–390.
- Bresnan, Joan: 1990, 'Levels of Representation in Locative Inversion', invited address presented at the 13<sup>th</sup> GLOW Colloquium at St. John's College, Cambridge University, on April 6, 1990; revised and duplicated, Department of Linguistics, Stanford University, California; further revised as Bresnan: 1994a.
- Bresnan, Joan: 1991, 'Locative Case vs. Locative Gender', in Laurel A. Sutton and Christopher Johnson with Ruth Shields (eds.), *Proceedings of the Seventeenth Annual Meeting* of the Berkeley Linguistics Society, pp. 53-68.
- Bresnan, Joan: 1994a, 'Locative Inversion and the Architecture of Universal Grammar', Language 70, to appear.
- Bresnan, Joan: 1994b, 'Category Mismatches', paper presented to the Plenary Session of the 25th Annual Conference on African Linguistics, Rutgers University, March 26, 1994.
- Bresnan, Joan and Jonni M. Kanerva: 1989, 'Locative Inversion in Chicheŵa: A Case Study of Factorization in Grammar', *Linguistic Inquiry* 20, 1-50.
- Bresnan, Joan and Jonni M. Kanerva: 1992, 'The Thematic Hierarchy and Locative Inversion in UG. A Reply to Paul Schachter's Comments', in Tim Stowell and Eric Wehrli (eds.), Syntax and Semantics 26: Syntax and the Lexicon, Academic Press, New York, pp. 111– 125.
- Bresnan, Joan and Sam A. Mchombo: 1986, 'Grammatical and Anaphoric Agreement', in A. M. Farley, P. T. Farley and K.-E. McCullough (eds.), *Papers from the Parasession on Pragmatics and Grammatical Theory*, CLS 22, Chicago, pp. 278–297.

- Bresnan, Joan and Sam A. Mchombo: 1987, 'Topic, Pronoun, and Agreement in Chicheŵa', Language 63, 741–782.
- Bresnan, Joan and Lioba Moshi: 1990, 'Object Asymmetries in Comparative Bantu Syntax', Linguistic Inquiry 21, 147-185.
- Butt, Miriam: forthcoming, 'Complex Predicate Scrambling' in Miriam Butt, Tracy Holloway King and Gillian Ramchand (eds.), Word Order in South Asian Languages, CSLI, Stanford University.
- Carstens, Vicki May: 1991, The Morphology and Syntax of Determiner Phrases in Kiswahili, Ph.D. dissertation, University of California, Los Angeles.
- Cho, Young-mee Yu and Peter Sells: forthcoming, 'A Lexical Account of Inflectional Suffixes in Korean', *Journal of East Asian Linguistics*.
- Chomsky, Noam: 1986, Barriers, MIT Press, Cambridge.
- Cole, Desmond T.: 1955, Introduction to Tswana Grammar, Longmans, Green, London and Capetown.
- Clark, Eve V. and Herbert H. Clark: 1979, 'When Nouns Surface as Verbs', Language 55, 767-811.
- Corbett, Greville G. and Alfred D. Mtenje: 1987, 'Gender Agreement in Chicheŵa', Studies in African Linguistics 18, 1–38.
- Dalgish, Gerard M.: 1976a, 'Locative NPs, Locative Suffixes and Grammatical Relations', in H. Thompson, K. Whistler, V. Edge, J. J. Jaeger, R. Javkin, M. Petruck, C. Smeall and R. D. Van Valin Jr. (eds.), Proceedings of the Second Annual Meeting of the Berkeley Linguistics Society, pp. 139-148.
- Dalgish, Gerard M.: 1976b, 'Passivizing Locatives in OluTsootso', Studies in the Linguistic Sciences 6, 57-68.
- Dalgish, Gerard M. and Gloria Sheintuch: 1976, 'On the Justification for Language-Specific Sub-Grammatical Relations', *Studies in the Linguistic Sciences* 6, 89–107.
- Demuth, Katherine and Mark Johnson: 1989, 'Interaction Between Discourse Functions and Agreement in Setawana', Journal of African Languages and Linguistics 11, 21-35.

Demuth, Katherine and Sheila Mmusi: 1992, 'Syntactic, Morphological, and Thematic Characteristics of Setswana Presentational Focus Constructions,' paper presented at the 23<sup>rd</sup> Annual Conference on African Linguistics, Michigan State University, East Lansing.

Di Sciullo, Anna-Maria and Edwin Williams: 1987, On the Definition of Word, MIT Press, Cambridge.

- Dixon, R. M. W.: 1977, 'Some Phonological Rules in Yidin'', Linguistic Inquiry 8, 1-34.
- Doke, Clement M.: 1929, 'The Problem of Word-Division in Bantu, With Special Reference to the Languages of Mashonaland', *Occasional Paper No. 2*, Department of Native Development, Southern Rhodesia.
- Doke, Clement M.: 1935, Bantu Linguistics Terminology, Longmans, Green, London.
- Doke, Clement M.: 1954, The Southern Bantu Languages, Oxford University Press, London.
- Doke, Clement M.: 1988, Textbook of Zulu Grammar, sixth edition, Maskew Miller Longman, Cape Town.
- Doke, Clement M. and S. M. Mofokeng: 1985, Textbook of Southern Sotho Grammar, second edition, Maskew Miller Longman, Cape Town.
- Fortune, G.: 1955, An Analytical Grammar of Shona, Longmans, London.
- Fortune, G.: 1985, Shona Grammatical Constructions: Parts 1 and 2, third edition, Mercury Press, Harare.
- Garrett, Andrew: 1990, 'Applicatives and Preposition Incorporation', in K. Dziwirek, P. Farrell and E. Mejías-Bikandi (ed.), *Grammatical Relations: A Cross-Theoretical Perspective*, CSLI, Stanford University, pp. 183-198.
- Givón, Talmy: 1972, Studies in ChiBemba and Bantu Grammar, Studies in African Linguistics3, supplement 3.
- Givón, Talmy: 1976, 'Topic, Pronoun, and Grammatical Agreement', in Charles N. Li (ed.), Subject and Topic, Academic Press, New York, pp. 149-188.

- Greenberg, Joseph H.: 1977, 'Niger-Congo Noun Class Markers: Prefixes, Suffixes, Both or Neither', *Studies in African Linguistics*, Supplement 7, 97-104.
- Greenberg, Joseph H.: 1978, 'How Does a Language Acquire Gender Markers?' in Joseph H. Greenberg (ed.), *Universals of Human Language*, vol. 3, Stanford University Press, Stanford, pp. 48–82.
- Guthrie, Malcolm: 1948, The Classification of Bantu Languages, Oxford University Press, London.
- Haider, Hubert and Martin Prinzhorn (eds.): 1986, Verb Second Phenomena in Germanic Languages, Foris Publications, Dordrecht.
- Heine, Bernd and Ulrike Claudi: 1986, On the Rise of Grammatical Categories. Some Examples from Maa, Dietrich Reimer Verlag, Berlin.
- Hodges, Kathryn S. and Susan U. Stucky: 1979, 'On the Inadequacy of a Grammatical Relation Referring Rule in Bantu', *Studies in the Linguistic Sciences* 9, 91-99.
- Homburger, L.: 1929, Les préfixes nominaux dans les parlers peul, haoussa et bantous, Institut d'ethnologie, Paris.
- Inkelas, Sharon: 1989, *Prosodic Constituency in the Lexicon*, Ph.D. dissertation, Stanford University.
- Inkelas, Sharon and Draga Zec (eds.): 1990, *The Phonology-Syntax Connection*, The University of Chicago Press, Chicago.
- Ishikawa, Akira: 1985, Complex Predicates and Lexical Operations in Japanese, Ph.D. dissertation, Stanford University.
- Kanerva, Jonni M.: 1987, 'Morphological Integrity and Syntax: The Evidence from Finnish Possessive Suffixes', *Language* 63, 498-521.

Kanerva, Jonni M: 1990, Focus and Phrasing in Chicheŵa Phonology, Garland, New York.

- King, Tracy: 1992, Configuring Topic and Focus in Russian, Ph.D. dissertation, Stanford University.
- Kinyalolo, Kasangati: 1991, Syntactic Dependencies and the SPEC-head Agreement Hypothesis in Kilega, Ph.D. dissertation, University of California, Los Angeles.
- Kiparsky, Paul: 1982, 'Lexical Phonology and Morphology', in S.-S. Yang (ed.), Linguistics in the Morning Calm, Hanshin, Seoul, pp. 3-91.
- Koopman, Hilda: 1984, The Syntax of Verbs: from Verb Movement Rules in the Kru Languages to Universal Grammar, Foris Publications, Dordrecht.
- Kroeger, Paul: 1993, Phrase Structure and Grammatical Relations in Tagalog, CSLI Publications, Stanford.
- Kuperus, Julianna and Mpunga wa Ilunga: 1990, *Locative Markers in Luba*, Annales Sciences Humaines 130, Tervuren, Musée royal de l'Afrique Centrale.
- Lapointe, Steven: 1979, 'A Lexical Analysis of the English Auxiliary Verb System', GLOT 2, 215-254.
- Lanham, L. W.: 1955, A Study of Gitonga of Inhambane, Witwatersrand University Press, Johannesburg.
- Lehmann, Christian: 1982, 'Universal and Typological Aspects of Agreement', in H. Seiler and F. J. Stachowiak, *Apprehension*, vol. II, Narr, Tübingen, pp. 201–267.
- Lehmann, Christian: 1988, 'On the Function of Agreement', in Michael Barlow and Charles Ferguson (eds.) Agreement in Natural Language: Approaches, Theories, Descriptions, CSLI, Stanford, pp. 55-65.
- Liberman, Mark and Richard Sproat: 1992, 'The Stress and Structure of Modified Noun Phrases in English', in Ivan A. Sag and Anna Szabolcsi (eds.), *Lexical Matters*, CSLI, Stanford University, pp. 131–181.
- Lieber, Rochelle: 1983, 'Argument Linking and Compounds in English', *Linguistic Inquiry* 14: 251-85.
- Lieber, Rochelle: 1988, 'Phrasal Compounds in English and the Morphology-Syntax Interface', in D. Brentari, G. Larson and L. MacLeod (eds.), *Papers from the Parasession on Agreement in Grammatical Theory*, CLS 24, Chicago, pp. 202–222.

- Lieber, Rochelle: 1992, Deconstructing Morphology: Word Formation In Syntactic Theory, The University of Chicago Press, Chicago.
- Manning, Christopher: 1992, 'Romance is so Complex', Report No. CSLI-92-168, CSLI, Stanford University.
- Matsumoto, Yo: 1992, On the Wordhood of Complex Predicates in Japanese, Ph.D. dissertation, Stanford University.
- Mchombo, Sam A.: 1978, A Critical Appraisal of the Place of Derivational Morphology within Transformational Grammar, Considered with Primary Reference to Chicheŵa and Swahili, Ph.D. dissertation, School of Oriental and African Studies, University of London.

Meinhof, Carl: 1906, Grundzuge einer vergleichenden Grammatik der Bantusprachen, Berlin.

Mohanan, K. P.: 1982, Lexical Phonology, Ph.D. dissertation, MIT.

Mohanan, K. P.: 1986, The Theory of Lexical Phonology, Reidel, Dordrecht.

- Mohanan, Tara: 1993, 'Case Alternation on Objects in Hindi', *South Asian Language Review* **3**, 1–30.
- Mohanan, Tara: forthcoming, 'Wordhood and Lexicality: Noun Incorporation in Hindi', Natural Language & Linguistic Theory.
- Mould, Martin: 1977, 'On the Productivity of Derivational Morphology and Lexical Representation: Manner Adverbs in Luganda', *Studies in African Linguistics*, Supplement 7, 175–183.
- Mtenje, Alfred D.: 1988, 'On Tone and Transfer in Chicheŵa Reduplication', *Linguistics* 16, 125-55.
- Mufwene, Salikoko: 1980, 'Bantu Class Prefixes: Inflectional or Derivational?', in Jody Kreiman and Almerindo E. Ojeda (eds.), CLS 16, Chicago, pp. 246-58.
- Myers, Scott P.: 1987, Tone and the Structure of Words in Shona, Ph.D. dissertation, The University of Massachusetts, Amherst.
- Nespor, Marina: 1985, 'The Phonological Word in Italian', in H. van der Hulst and N. Smith (eds.), Advances in Nonlinear Phonology, Foris Publications, Dordrecht, pp. 193–204.
- Nespor, Marina and Irene Vogel: 1986, Prosodic Phonology, Foris Publications, Dordrecht.
- Netter, Klaus and Ursula Kärcher, 1986, Documentation of the German Grammar, ESPRIT ACORD project 393, Deliverable Task 1.4., Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.
- Netter, Klaus: 1988, Syntactic Aspects of LFG-based Dialogue Parsing, ESPRIT ACORD project 393, Deliverable Task 2.7'(a), Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.

Odden, David: 1992, 'C-command or Edges in Makonde', Phonology 7, 163-169.

- Orr, Gregory John and Carol Myers Scotton: 1980, Learning Chicheŵa, Books 1 and 2, A Peace Corps Language Course Prepared by the African Studies Center, Michigan State University, East Lansing.
- Perez (Harford), Carolyn: 1983, 'Locative Pseudo-Subjects in Shona', Journal of African Languages and Linguistics 5, 131-155.
- Poser, William J.: 1990, 'Word-internal Phrase Boundary in Japanese', in Sharon Inkelas and Draga Zec (eds.), *The Phonology-Syntax Connection*, The University of Chicago Press, Chicago, pp. 279–287.
- Poser, William J.: 1991, 'Japanese Periphrastic Verbs and Noun Incorporation', unpublished manuscript, Stanford University.
- Postal, Paul: 1969, 'Anaphoric Islands', in Robert I. Binnick, Alice Davison, Georgia M. Green and Jerry L. Morgan (eds.), CLS 5, Chicago.
- Sadock, Jerry M.: 1980, 'Noun Incorporation in Greenlandic: A Case of Syntactic Word Formation', *Language* 57, 300-319.
- Sadock, Jerry M.: 1991, Autolexical Syntax: A Theory of Parallel Grammatical Representation, The University of Chicago Press, Chicago.
- Selkirk, Elisabeth O.: 1982, The Syntax of Words, MIT Press, Cambridge.
- Selkirk, Elisabeth O.: 1984, Phonology and Syntax: The Relation between Sound and Structure, MIT Press, Cambridge.

Sells, Peter: 1992, 'Japanese and Korean Morphology from a Lexical Perspective', *Linguistic Inquiry*, to appear.

Sells, Peter and Young-mee Yu Cho: 1991, 'On the Distribution of X<sup>0</sup> Elements in Korean', Proceedings of the Santa Cruz Workshop on Korean Syntax and Semantics, UCSC.

Shibatani, M. and T. Kageyama: 1988, 'Word Formation in a Modular Theory of Grammar: A Case of Post-Syntactic Compounds in Japanese', *Language* 64, 455–84.

Simpson, Jane: 1983, Aspects of Warlpiri Morphology and Syntax, Ph.D. dissertation, MIT.

Simpson, Jane: 1991, Warlpiri Morpho-Syntax: A Lexicalist Approach, Kluwer, Dordrecht.

Speas, Margaret: 1991, 'Functional Heads and Inflectional Morphemes', *The Linguistic Review* 8, 389-417.

Spencer, Andrew: 1988, 'Bracketing Paradoxes and the English Lexicon', Language 64, 663–82.

Spencer, Andrew: 1991, Morphological Theory, Basil Blackwell, Oxford, pp. 413-417.

Sproat, Richard: 1985, On Deriving the Lexicon, Ph.D. dissertation, MIT.

Stowell, Timothy: 1981, Origins of Phrase Structure, Ph.D. dissertation, MIT.

Stucky, Susan U.: 1976, 'Locatives as Objects in Tshiluba: a Function of Transitivity', Studies in the Linguistic Sciences 6, 174-202; revised as Stucky: 1978.

- Stucky, Susan U.: 1978, 'Locative Phrases and Alternative Concord in Tshiluba', Studies in African Linguistics 9, 107-119.
- Tabor, Whitney: 1992, 'Auxiliary Coalescence in Chicheŵa: Mismatch Not Required', in Jonathan Mead (ed.), *Proceedings of the 11th West Coast Conference on Formal Linguistics* (WCCFL), CSLI, Stanford.
- Travis, Lisa: 1984, Parameters and Effects of Word Order Variation, Ph.D. dissertation, MIT.
- Trithart, Lee: 1977, 'Locatives', in E. Byarushengo, A. Duranti and L. Hyman (eds.), Haya Grammatical Structure, Southern California Occasional Papers m Linguistics No. 6, University of Southern California, Los Angeles, pp. 89–98.

Ward, Gregory, Richard Sproat and Gail McKoon: 1991, 'A Pragmatic Analysis of So-called Anaphoric Islands', *Language* 67, 439–474.

Welmers, William E.: 1973, African Language Structures, University of California Press, Berkeley.

Received 25 January 1993 Revised 14 September 1993

Bresnan CSLI, Ventura Hall Stanford University Stanford, CA 94305 bresnan@csli.stanford.edu

Mchombo Department of Linguistics University of California, Berkeley 2337 Dwinelle Hall Berkeley, CA 94720 mchombo@garnet.berkeley.edu