

## **DP in Bantu and Romance\***

Vicki Carstens

University of Missouri-Columbia

[carstensv@missouri.edu](mailto:carstensv@missouri.edu)

The paper considers several aspects of Bantu and Romance DPs, including (i) Noun Class and grammatical gender; (ii) apparently derivational properties of the two; (iii) ordering among nouns and their modifiers, and (iv) concord in DPs. Several conclusions are argued for. First, Bantu Class is a gender system like that of Romance, with gender-specific Spell-Out of number features. Second, despite some surface evidence to the contrary, gender/Class is an uninterpretable feature, without derivational functions. Third, DPs of the two languages share a common architecture; and fourth, in both families concord is the result of the Agree relation. Thus many properties of nouns and DPs are common to both language groups, as the hypothesis of UG leads us to expect.

### 1. Introduction.

Languages of the Bantu and Romance families exhibit striking similarities, including several points of correspondence within the DP: members of both groups partition nouns into genders or classes; categories within DP agree with the head noun in this feature and in number; and movement places both Bantu and Romance nouns to the left of their base positions. This paper compares and contrasts these aspects of DP morpho-syntax in the two families, delineating both underlying unities and points of true divergence.

I begin in 2.1 with an analysis of Bantu Noun Class as a gender system very much like that of Romance. The remainder of section 2 explores some apparently derivational functions of Noun Class that lack precise parallels in Romance. I show that, contrary to appearances, neither Noun Classes nor their prefixes are derivational; rather there are derivational zero morphemes in Bantu with their own gender specifications. These are kin to the larger, phrase-level null element *pro* found in both Bantu and Romance, and like *pro* they are subject to a recoverability or identification requirement. Bantu has a greater number and diversity of null elements than Romance does because its larger number of genders enables them to be identified unambiguously and hence to be recoverable. This is a modest difference, consistent with the assumption that the underlying grammatical principles at work are very much the same.

Section 3 compares the syntax of Romance and Bantu DPs. I argue that the architecture is shared, and that both groups of languages exhibit noun raising, though they differ in the landing sites of nouns. I consider and reject the anti-symmetric approach to post-nominal modifier order presented in Cinque (2005) in favor of symmetric base-generation of modifiers.

Section 4 shows that in both groups of languages the Agree relation (cf. Chomsky 2000; 2001) can readily account for the agreement of categories within the DP with the head noun (often labeled concord).

Section 5 concludes that all aspects of DPs in the two languages are indeed highly similar – a welcome finding, supportive of the theory of Universal Grammar.

For the sake of brevity I rely as much as possible on representative examples from just two languages: Swahili for illustrations of Bantu phenomena, and Spanish for illustrations of Romance phenomena. Because the paper deals with broad questions this is generally sufficient.

## 2. Words and morphemes

### 2.1 Noun Class as grammatical gender

As is well-known, Romance languages generally partition nouns into two grammatical genders, masculine and feminine. Most nouns bear a suffixal word-marker whose shape correlates fairly consistently with the gender of the noun (see 1).<sup>1</sup> In Spanish, a plural suffix may follow it, as in (2). Nouns denoting humans are distributed among the two genders on the basis of the sex of their referents (see 3).

- (1) a. cas-a  
house-F  
'house' [Spanish]
- b. libr-o  
book-M  
'book'
- (2) a. cas-a-s  
house-F-PL  
'houses'
- b. libr-o-s  
book-M-PL  
'books'
- (3) a. chic-a  
child-F  
'girl'
- b. hij-o  
offspring-M  
'son'

Bantu nouns are similarly partitioned, though there are several points of contrast. First, natural or sex gender is not a factor in determining the Classes of nouns; rather, nouns

denoting humans or, in many Bantu languages, animates, constitute a single Class. Second, Bantu Noun Classes are indicated by prefixes; and third, they greatly outnumber the Romance genders, especially as the Classes are conceptualized in the Bantuist linguistic tradition. I provide representative examples from Swahili in (4). The numbers of the Noun Classes are traditional ones, based on Meinhoff's reconstructions for proto-Bantu. Classes 12 and 13 of Proto-Bantu are absent in Swahili;<sup>2</sup> on locative Classes 16-18, see section 2.4.

(4)	Noun Classes: <i>Swahili</i>	Class	Example	Gloss
		1	<i>m-tu</i>	person
		2	<i>wa-tu</i>	people
		3	<i>m-ti</i>	tree
		4	<i>mi-ti</i>	trees
		5	<i>gari</i>	car
		6	<i>ma-gari</i>	cars
		7	<i>ki-atu</i>	shoe
		8	<i>vi-atu</i>	shoes
		9	<i>n-yumba</i>	house
		10	<i>n-yumba</i>	houses
		11	<i>u-bao</i>	board
		14	<i>u-kweli</i>	truth
		15	<i>ku-soma</i>	to read; reading
		16	N+loc. suffix	specific place
		17	N+loc. suffix	general place
		18	N+loc. suffix	inside place

It is clear that pairs of prefixes for many Classes function as singular/plural morphology for particular groups of stems. For example every noun stem in Swahili which takes the Class 1 singular prefix also takes the Class 2 prefix in the plural (see 5); every noun which takes the Class 7 singular prefix takes the Class 8 prefix in the plural (see 6); and so on, through Class 10. Nouns in Class 11 are exceptional to a very modest degree, in that their plurals are in Class 10, like the plurals of Class 9 nouns.

- (5)
- |    |   |           |
|----|---|-----------|
| a. | mtoto/watoto<br>1child/2child<br>'child/ren'  | [Swahili] |
| b. | mjinga/wajinga<br>1fool/2fool<br>'fool/s'     |           |
| c. | msichana/wasichana<br>1girl/2girl<br>'girl/s' |           |

- (6) a. kiatu/viatu  
7shoe/8shoe  
'shoe/s'
- b. kikombe/vikombe  
7cup/8cup  
'cup/s'
- c. kikapu/vikapu  
7basket/8basket  
'basket/s'

It is a weakness of the traditional system depicted in (4) that the regular relationships among pairs of Classes is not formally reflected; though it has the virtue of familiarity and conserves some space in glosses, the system is theoretically un insightful. This is demonstrated in (7) through its hypothetical extension to a Romance language; I replace the gender and number designations of the Spanish nouns from (1) and (2) with Bantu-style Class numbers. A feminine noun like *casa* - 'house' or *chica* - 'girl' is analyzed as belonging to two different classes, 1 and 2, depending on whether it is singular or plural; similarly for *libro* - 'book' or *hijo* - 'son', whose singulars are Class 3 and plurals Class 4. We lose the insight that there are two groups of stems, not four.

- |     |    |                             |    |                               |           |
|-----|----|-----------------------------|----|-------------------------------|-----------|
| (7) | a. | casa<br>house1<br>'house'   | b. | chica<br>girl1<br>'girl'      | [Spanish] |
|     | c. | casas<br>house2<br>'houses' | d. | chicas<br>girl2<br>'girls'    |           |
| (8) | a. | libro<br>book3<br>'book'    | b. | hijo<br>offspring3<br>'son'   |           |
|     | c. | libros<br>book4<br>'books'  | d. | hijos<br>offspring4<br>'sons' |           |

Because plurality has a consistent realization as the suffix *-s* in the Spanish data, the approach is transparently misguided. In Bantu, on the other hand, the realization of plural (and singular) varies by gender. But this is a difference in surface detail, not in underlying categories or concepts.<sup>3</sup>

I accordingly follow Carstens (1991) in analyzing Noun Class as a gender system, along the lines of Romance (see also Corbett 1991). This proposal reduces Noun Classes 1-11 to six genders, as shown in (9).<sup>4,5</sup> Prefixes are gender-specific number morphology, which I assume to be added to nouns by late Spell-Out rules (cf. Halle & Marantz 1993).<sup>6</sup>

The approach is neutral on the question of phrasal projections for inflectional features; I will argue in 2.2.1 that no such projection is motivated for gender; in contrast, interpretable number features may head functional projections (see section 3).

(9) Bantu Genders (Carstens 1991)

Gender A: stems of Classes 1/2  
 Gender B: stems of Classes 3/4  
 Gender C: stems of Classes 5/6  
 Gender D: stems of Classes 7/8  
 Gender E: stems of Classes 9/10  
 Gender F: stems of Classes 11/10

(10) *Sample Spell-Out rules yielding Swahili Noun Class Prefixes<sup>7</sup>*

[Singular] < -- > /m- / /\_\_N  
 Gender A

[Singular] < -- > /ki- / /\_\_N  
 Gender D

[Plural] < -- > /vi- / /\_\_N  
 Gender D

[Plural] < -- > /wa- / /\_\_N  
 Gender A

In contrast, a Spell-Out rule for singular number in Spanish is superfluous, since it would consistently add Ø; and since the realization of plural is not gender-specific in Spanish, a single Spell-Out rule suffices for all regular plurals:

(11) [Plural] < -- > /-s / /N\_\_ *Spell-Out for the Spanish plural suffix*

Summing up, I have sketched out a treatment of Noun Class in line with the following general conclusions:

- (12) a. Bantu and Romance both have grammatical gender.  
 b. Bantu has a greater number of genders than Romance.  
 c. Bantu expresses number in gender-particular prefixes, while Romance concatenates markers of gender and number as suffixes.  
 d. Animacy or humanness has a gender correlation in Bantu languages; biological sexes have such correlates in the genders of Romance.

The differences between Romance gender and Bantu Noun Class are fairly trivial under this approach, and this is a desirable conclusion, consistent with the view that deep unities underlie surface variation among languages. But for the analysis of Noun Class as

gender to succeed, Classes 14-18 must be given a satisfactory account, as must a set of derivational phenomena associated with certain Classes. I turn to these matters next.

## 2.2 Noun Class and derivation

### 2.2.1 Representing Class

Obscuring the simplicity of the Bantu system and the unity of Bantu and Romance are some much-noted alternations in which Noun Class prefixes seem to play a derivational role. Observe that replacing a noun's usual Class prefixes with those of Classes 7 and 8 yields diminutive meanings in Swahili; switching to those of Classes 5 and 6 produces augmentative meanings (augmentatives are also formed in Classes 3/4, not exemplified here):

- |      |    |   |    |   |           |
|------|----|---|----|---|-----------|
| (13) | a. | msumari/misumari<br>3nail/4nail<br>'nail/s'                 | b. | kisumari/visumari<br>7nail/8nail<br>'little nail/s' | [Swahili] |
|      | c. | sumari/masumari<br>5nail/6nail<br>'big (ugly/nasty) nail/s' |    |   |           |

When the prefixes are replaced in this way, agreement with the noun also shifts, reflecting the Class of the replacement prefix.<sup>8</sup>

- |      |    |  |           |
|------|----|--|-----------|
| (14) | a. | msumari mmoja u-me-anguka<br>3nail 3one 3agr-PERF-fall<br>'one nail has fallen'            | [Swahili] |
|      | b. | kisumari kimoja ki-me-anguka<br>7nail 7one 7agr- PERF-fall<br>'one little nail has fallen' |           |
|      | c. | masumari mawili ya-me-anguka<br>6nail 6big 6agr- PERF-fall<br>'two big nails have fallen'  |           |

Two conclusions have been drawn from this state of affairs. First, many researchers have argued from such facts as (13) that Noun Class is entirely a property of the Noun Class prefixes, which are heads, selecting N(P)s (cf. Sproat 1985, Myers 1987, Bresnan & Mchombo 1995). Thus the representation of a noun is as shown in (15), in these works.



Second, alternations like (14) are taken as evidence that Noun Class is at least partially derivational in nature (Mufwene 1980, Myers 1987, and Ferrari 2005).

In what follows I will argue that there are empirical reasons to reject both of these views.

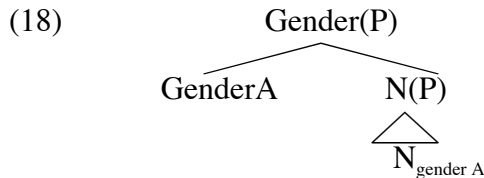
Consider first the status of stems with respect to Class information. Although they are treated as unmarked for Class in Sproat (1985), Myers (1987), Bresnan & Mchombo (1995), this cannot be correct. We must assume that stems are specified for Class (in the framework adopted here, for gender) to account for combinatorial restrictions like (16) and for the default pattern of singular and plural formation for each noun. Thus something like (9) seems to be unavoidable.

- (16) a. \*n-tu  
C9-person  
'person'
- b. \*mi-atu  
C3-shoe  
'shoe'

One might solve this problem by specifying the stems for gender while still maintaining the analysis of Noun Class prefixes as selecting heads. But this creates an undesirably redundant system, since each member of a singular-plural prefix pair must be specified for identical selectional features (see 17).

- (17) a. *m-* [ $\_ + \alpha_{\text{Gender A}}$ ] denotes singular  
*wa-* [ $\_ + \alpha_{\text{Gender A}}$ ] denotes plural
- b. *m-* [ $\_ + \alpha_{\text{Gender B}}$ ] denotes singular  
*mi-* [ $\_ + \alpha_{\text{Gender B}}$ ] denotes plural
- c.  $\emptyset$  [ $\_ + \alpha_{\text{Gender C}}$ ] denotes singular  
*ma-* [ $\_ + \alpha_{\text{Gender C}}$ ] denotes plural
- d. *ki-* [ $\_ + \alpha_{\text{Gender D}}$ ] denotes singular  
*vi-* [ $\_ + \alpha_{\text{Gender D}}$ ] denotes singular  
etc.

An alternative to considering the specific prefixes to be selecting heads might be to suppose that gender as a general category is a head selecting nominal complements, or perhaps that each of the Bantu genders proposed in (9) is a head selecting stems with matching gender specifications (see 18). The prefixes themselves would spell out singular or plural on a gender-specific basis, at the point that number features are added, as shown in (10).



This proposal too has serious drawbacks. Chomsky (1995) proposes that uninterpretable features must be deleted from the representations sent to the semantic interface, and that it is therefore unlikely that they head syntactic categories, introducing structure which must subsequently be eliminated. Though the specific concern of his proposal is agreement features, the line of reasoning extends naturally to gender, arguing against a syntactic representation like (18), if gender is an uninterpretable  $\phi$ -feature like agreement features are.

There are good reasons to think that gender is in fact uninterpretable. Despite some pockets of semantic unity in the groups of nouns assigned to one gender or another, membership in a grammatical gender is recognized to be, on balance, arbitrary. There is no semantic reason why, for example, the Swahili noun *kikapu* - ‘basket’ should belong to Gender D, while the noun *ndoo* - ‘bucket’ is Gender E; or why the Italian *patata* - ‘potato’ should be feminine while *pomodoro* - ‘tomato’ is masculine.<sup>9</sup>

Even where there is some semantic consistency to gender, as in the assignment of animates to Swahili Gender A or females to Romance feminine gender, this does not entail that the gender itself has any semantic content; rather, some semantic properties of the nouns seem to serve as sorting criteria, similar to what is involved in a household decision to keep plates on one shelf and glasses on another. No one considers that essential properties of plates and glasses are intrinsic to the shelves on which we choose to store them; similarly, threads of semantic unity in grammatical genders do not indicate that the genders themselves have semantic content.

These considerations argue in favor of viewing gender as an uninterpretable feature of nouns, and hence against adopting (18).

It has often been claimed that consistent meanings like diminutive and augmentative are properties of specific Bantu Noun Classes and/or their prefixes. If this should turn out to be true, (18) might yet be a possibility. I will argue in 2.2.2 that these meanings are instead properties of zero-morphemes of particular genders. I conclude that gender is not a selecting head, and should not have a projection of its own.



### 2.2.2 The zero-affixation analysis

In this section I will provide two arguments that diminutives and augmentatives are formed by a zero-affixation process distinct from the attachment of Noun Class prefixes, contra Sproat (1985), Myers (1987), Bresnan & Mchombo (1995), Mufwene (1980) and Ferrari (2005).

The first argument is based on a minimum word size requirement and an epenthetic noun prefix *ji-*. Lexical categories in Swahili must consist of at least two moras:

- (19) Swahili Two Mora Constraint:  $*[N \mu^n]$ ,  $n < 2$  (cf. Carstens 1991; Park 1997; Njogu 1994)

Noun class prefixes enable many monomoraic stems to meet this requirement:

- (20) a. mtu/watu                      b. kiti/viti                      [Swahili]  
1person/2person                      7chair/8chair  
'person/people'                      'chair/s'
- c. mti/miti  
3tree/4tree  
'tree/s'

Class 5 (= the singular of Gender C in 9) generally lacks any prefix (see 21). Monomoraic stems in this Class surface bearing the prefix *ji-* (see 22), absent on polymoraic Class 5 nouns. I accordingly analyze *ji-* as an epenthetic element with the function of “rescuing” stems which would otherwise violate the Two Mora Constraint:

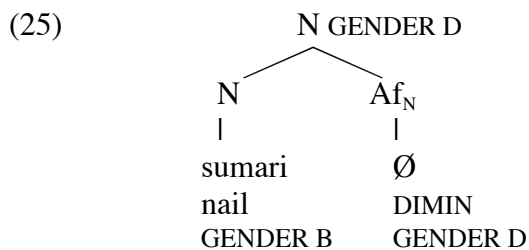
- (21) a. bega/mabega                      [Swahili]  
5shoulder/6shoulder  
'shoulder/s'
- b. tawi/matawi  
5branch/6branch  
'branch/es'
- (22) a. *ji*cho/macho  
5eye/6eye  
'eye/s'
- b. *ji*we/mawe  
5stone/6stone  
'stone/s'

- (23) *ji*-insertion: Prefix *ji-* to a noun which fails to satisfy the Two Mora Constraint.<sup>10</sup>

*Ji*-epenthesis interacts with derivation in revealing ways. Recall that diminutive nouns belong to Gender D, and thus bear *ki/vi* prefixes; augmentatives belong either to Gender C (Classes 5 and 6) or Gender B (Classes 3 and 4) of which all but the singular Class 5 have overt Noun Class Prefixes. Yet any monomoraic stem made diminutive or augmentative must bear *ji-* as a kind of “inner” prefix, despite the presence of the Noun Class prefixes of its derived Class:

- (24) a.      *kijicho/vijicho*                      b. \*      *kicho/vicho*                                      [Swahili]  
              7eye/8eye                                      7eye/8eye  
              ‘envy; jealous glance/s’                      ‘envy; jealous glance/s’  
              (lit: little eye/s)
- b.      *mjicho/mijicho*                      d. \*      *mcho/micho*  
              3/eye/4eye                                      3eye/4eye  
              ‘big eye/s’    ‘big eye/s’

As Carstens (1991) points out, two assumptions allow us to make sense of this distribution of *ji-*: (a) the Two Mora Constraint and *ji*-epenthesis apply in derived environments;<sup>11</sup> and (b) the derivational processes forming diminutives and augmentatives are zero-affixation processes distinct from the attachment of Noun Class prefixes. I assume that they involve silent morphemes with their own semantics and gender specifications, as shown in (25); since overt derivational affixes are suffixes, in Swahili, I treat the zero-affixes as suffixes also.<sup>12 13</sup> The diminutive affix selects the noun stem and is therefore head of the derived word, determining its category and gender specification.



Under these assumptions, a noun stem made diminutive or augmentative is subject to the Two Mora Constraint because zero-affixation creates a derived environment. On the other hand, the monomoraic roots in (21), which bear no derivational affixes, are assessed for compliance with the Two Mora Constraint only when the addition of singular or plural Noun Class prefixes creates a derived environment. The Two Mora Constraint therefore applies only to the prefixed form in such cases, and is satisfied without *ji*-insertion.

An analysis equating diminutive and augmentative formation with the prefixes of the derived forms would be hard pressed to account for these facts, since the prefixes *ki-*, *vi-*, *m-* and *mi-* and *ma-* do not otherwise require that the stems they attach to meet the Two Mora Constraint (see 20b,c).<sup>14</sup> One could only maintain the approach by supposing that

diminutives and augmentatives are formed, rather redundantly, by means of pairs of prefixes that happen to be identical and homophonous to existing singular and plural prefixes, but which differ from them in having the semantics [diminutive singular] and [diminutive plural], [augmentative singular] and [augmentative plural]; these prefixes also happen, exceptionally, to require that the stems they attach to meet the constraint. This is neither economical nor advantageous in any other way that I can think of, so I reject it in favor of the simpler and more elegant zero-affixation account.

I turn to the second argument that diminutive and augmentative formation processes are distinct from their associated genders and prefixes. Though ordinary nouns of Classes 7/8 bear *ki-/vi-* prefixes like diminutive nouns, they do not necessarily denote small things (see 26 a,b), nor must those of Classes 5/6 denote large things (see 26c,d) despite its apparent role in augmentative formation. It follows that the diminutive and augmentative semantics do not reside in the relevant genders or their associated prefixes, contra Mufwene (1980) and Ferrari (2005). The analysis in (25) is consistent with this state of affairs.

- |      |    |  |    |                                     |           |
|------|----|--|----|-------------------------------------|-----------|
| (26) | a. | Kiingereza<br>7English<br>'English language' | b. | kiapo<br>7oath<br>'oath' (deverbal) | [Swahili] |
|      | c. | maneno<br>6word<br>'words'                   | d. | matawi<br>6branch<br>'branches'     |           |

In fact, ordinary nouns of Classes 7/8 are not even ambiguous between neutral and diminutive readings, nor are those of Classes 5/6 ambiguous between neutral and augmentative readings (see 27).<sup>15</sup>

- |      |    |  |    |   |           |
|------|----|--|----|---|-----------|
| (27) | a. | kitabu/vitabu<br>7book/8book<br>'book/s'<br>*'little book/s' | b. | gari/magari<br>5car/6car<br>'car/s'<br>*'big car/s' | [Swahili] |
|------|----|--|----|---|-----------|

I follow Carstens (1991) in relating these facts to independently motivated licensing requirements on zero-morphemes (cf. Pesetsky 1990), which can in turn be related to general recoverability requirements. The augmentative and diminutive morphemes are subject to (28):<sup>16</sup>

- (28) Zero-morpheme licensing principle: Zero-morphemes must be identified.

Pesetsky (1990) argues that inflection and certain derivational morphemes can identify zero-morphemes in English; Carstens (1991) proposes that the requirement is satisfied for Bantu zero-morphemes by their introduction of a gender specification that contrasts with that of the root, as in (29a).<sup>17</sup>

- (29) a. *ok* [sumari<sub>Gender B</sub>] Ø<sub>Gender D</sub>] - ‘little nail’  
 b. \* [tabu<sub>Gender D</sub>] Ø<sub>Gender D</sub>] - \*‘little book’
- c. Zero-morpheme identification condition: a zero-morpheme is identified if its features are unambiguously represented on the derived word.

I will propose in section 2.5 that this requirement is kin to the well-known licensing requirement on *pro*. In contrast, is not at all clear how the facts in (26) and (27) can be accounted for under the assumption that Noun Classes or their prefixes have a derivational function.

## 2.3 Discussion

I have argued that Bantu Noun Class prefixes are gender-specific number morphology, and that derivational properties of the Classes and their prefixes are apparent rather than real. Diminutives and augmentatives in Bantu are zero-affixation processes, where the zero-affixes have gender specifications.

The Swahili diminutive morpheme differs from those found in Romance languages in that it is silent, and in that Romance diminutive morphemes are unspecified for gender. Under the morphological percolation conventions of Selkirk (1982), a derived word inherits the features of its head and any non-clashing features of its non-head. Thus absent a competing gender feature in the affix, a Spanish noun made diminutive inherits the gender of its non-head member, the stem (see 30).<sup>18,19</sup>

- (30) a. el carr-o  
the car (masc)  
'the car'
- b. el carr-it-o  
the car-DIMIN (masc)  
'the little car'
- c. la caj-a  
the box (fem)  
'the box'
- d. la caj-it-a  
the box-DIMIN (fem)  
'the little box'
- e.
- N [ $\alpha$  gender]
- N                  DIMIN
- [ $\alpha$  gender]

Bantu diminutives are more like those of German, which are always neuter (see 31), or Faroese, where diminutives are uniformly masculine (32). We can assume that the German diminutive morpheme has a neuter specification which is inherited by the derived noun, while that of Faroese has a feminine feature. (32) illustrates schematically diminutives of Faroese, German, and Bantu.

- (31) a. die maus  
the mouse (fem)  
'the mouse'
- b. das maus-chen [German]  
the mouse-DIMIN (neuter)  
'the little mouse'

- (32) c. der hase  
the hare (masc)  
'the hare'
- d. das has-chen  
the hare-DIMIN (neuter)  
'the little hare'
- (33) a. ketta  
cat (fem)  
'cat'
- b. kett-lingur [Faroese]  
cat- DIMIN (masc)  
'kitten'
- (33) 

```
graph TD
    A["N [α gender]"] --- B["N [β gender]"]
    A --- C["DIMIN [α gender]"]
    B --- D[ketta]
    C --- E[lingur]
```
- Bantu, German, and Faroese diminutives*

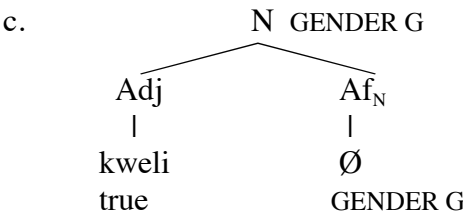
The Bantu pattern of diminutive formation thus differs from that of Romance in several ways. But under this line of analysis for derivational phenomena associated with Bantu Noun Classes, some otherwise exotic-seeming functions of Class dissolve, leaving us with an understanding of Class as a simple gender system that differs from that of Romance only in minor details.<sup>20</sup>

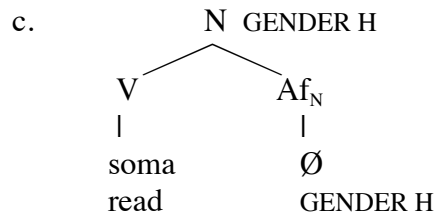
My approach to Bantu diminutives and augmentatives can be extended to abstract noun formation (34) and gerund formation (35), though these processes change categories; we need only assume that the zero morpheme head of the derived noun contributes a nominal categorial feature, as many derivational morphemes do (cf. *destroy/destruction*; *happy/happiness*, etc.). Abstract nouns and gerunds have nominal morphology and agreement paradigms distinct from other Classes, so I posit two additional genders for them, Genders G and H.

- (34) a. kweli  
true  
'true'
- b. u-kweli  
14-true  
'truth'
- [Swahili]
- c.
- ```

      N GENDER G
     /  \
  Adj    AfN
   |      |
kweli    Ø
true     GENDER G

```
- (35) a. soma  
read<sub>V</sub>  
'read'
- b. ku-soma  
15-read<sub>N</sub>  
'reading'





There are in Romance languages some correlations of semantics and gender that Ferrari (2005) cites in evidence of her claim that gender is semantic and derivational. As previously noted, biological gender of the referent correlates with grammatical gender in Romance nouns referring to humans:

- (36) a. la niña the.fem child.fem 'the girl' b. el niño the.masc boy.masc. 'the boy' [Spanish]
- c. la señora the.fem woman.fem 'the woman' d. el señor the.masc man.masc 'the man'

Ferrari (2005) describes a number of additional correlations of grammatical gender and semantic feature. There are a count vs. mass association for masculine vs. neuter gender in Ripano, and several patterns of semantic correlation to gender choice in Cantabrian (Ferrari 2005: 39-44), among them male/female, dark/light, coarse/smooth, vertical/horizontal, narrow/wide. Ferrari also provides examples of several systematic gender-meaning correspondences in Italian, which I reproduce below:<sup>21</sup>

- (37) a. ferramento 'iron tool' -ferramenta 'hardware' [count/mass]  
 b. grano 'grain' -granola 'the sifted whole wheat grains' [count/mass]  
 c. raccolto 'harvest/crop' - raccolta 'harvesting, collection' [+/- collective]  
 d. scarico 'unloading of a weapon' - scarica 'a volley of rifle-fire' [+/-collective]  
 e. taglio 'cut' ~ taglia 'ransom, tally, body size' [concrete/abstract]  
 f. rancio 'ration' ~ rancia 'the distribution of the 'ration' [concrete/abstract]  
 g. mestolo 'small ladle' ~ mestola 'big ladle' [small/big]  
 h. buco 'small hole' ~ buca 'pit or big hole' [small/big]  
 i. pozzo 'well' ~ pozza 'puddle' [small/big]  
 j. granito 'granite'~granita 'grated-ice drink' [inanimate- animate (food)]  
 k. ciuccio 'pacifier'~ciuccia 'breast' [inanimate –animate (body part)]  
 l. melo 'apple tree' ~ mela 'apple' [inanimate – animate (food)]  
 m. lucciollo 'sequin' - lucciola 'fire-fly' [inanimate-animate (insect)]  
 n. girello 'small disk' - girella 'spinning wheel' [small/big]

- o. ninfeo ‘temple of pagan gods’ - ninfea ‘a kind of plant’ [+/-animate]
- p. terrazzo ‘balcony’ ~ terrazza ‘terrace’ [small/big]

The co-existence of different strands of meaning within the same gender (for masculine: count or male or small or concrete or inanimate) presents a challenge to analyzing genders as contributing any meaning. Ferrari addresses this by proposing that semantic properties of the gender interact with those of the root to yield the correct meaning.

I believe there is a simple and very plausible interpretation of how root and gender interact that is consistent with the status of gender as an uninterpretable inflectional feature. An epicene noun (borrowing Ferrari’s term) can be viewed as one whose lexical meaning contains an open variable (underspecification, represented as  $\emptyset$  in 38 to avoid confusion with uninterpretability, often represented  $u$ ). Each such variable ranges over a pair of oppositions out of a fixed and finite set: male/female; mass/count; small/large; concrete/abstract; and so forth on a language-particular basis. The variability of such nouns is grammatically encoded in underspecification for gender (also represented as  $\emptyset$ ). A set of redundancy rules maps the nouns to genders, according to the value given to the semantic variable upon lexical selection:

(38) Redundancy rules assigning gender to epicene nouns

| <i>Root</i>                         | <i>Gender mapping rule</i>                     |
|-------------------------------------|------------------------------------------------|
| $N_{\emptyset_{\text{sex}}}$        | $\rightarrow$ feminine if female <sup>22</sup> |
| $\emptyset_{\text{gender}}$         | $\rightarrow$ masculine if male                |
| $N_{\emptyset_{\text{count/mass}}}$ | $\rightarrow$ masculine if count               |
| $\emptyset_{\text{gender}}$         | feminine if mass                               |
| $N_{\emptyset_{\text{size}}}$       | $\rightarrow$ masculine if small               |
| $\emptyset_{\text{gender}}$         | feminine if large                              |
| etc.                                |                                                |

Under this approach, the genders themselves contribute no meaning. It is therefore quite natural that the collection of semantic features in (38) that map to a given gender is an arbitrary set, just like the assignment of ordinary nouns to genders is a semantically arbitrary assignment.

Before leaving this topic, let me summarize the significant contrasts between Romance epicene nouns on the one hand and Bantu formation of diminutives, augmentatives, abstract nouns and gerunds on the other. Romance epicene nouns are a specific, finite set of nouns. While common themes are observable, the semantics involved in gender choice are determined by the roots themselves. In contrast, the Bantu processes are regular, productive derivational processes in which specific, predictable meanings can be linked to specific affixes. The analyses I have provided reflect these differences.

## 2.4 Locatives.

A final aspect of Bantu Noun Class that distinguishes it from Romance gender is the phenomenon of Locative Noun Classes. Many Bantu languages have three locative Classes, with characteristic agreement and Noun Class prefixes. I illustrate with Chichewa:

- (39) a. Mu-nyumba mu-ku-nunkha. [Chichewa]  
18-9house 18agr-ASP-stink-FV  
'Inside the house stinks'
- b. Ku-nyumba 'ku ndi ku-tali  
17-9-lake 17DEM COP 17AGR-far  
'That house and its environs are far away'
- c. Pa-nyumba pa-ku-on-ek-a ngati pa-ku-psy-a  
16-9house 16 AGR-ASP-see-STAT-FV like 16AGR-ASP-burn- FV  
'The house and surrounding yard look like they're burning'

I present here a sketch of the properties of the locative classes, and show that while they have no counterpart in the system of Romance genders, their syntax crucially involves sub-phrasal gaps that occur in Romance as well. In both language families the gaps can be licensed by "rich" agreement, indicating that the grammatical principles underlying the two systems are truly the same.

The structure of locative phrases in Bantu is revealed by the agreement and interpretation of modifiers within them. In many languages, including Chichewa and Shona, modifiers of locativized nouns can agree in either the locative Class or the Class of the noun stem. Though Myers (1987) asserts that there are no semantic correlates to this choice, Carstens (1997) shows otherwise. Items bearing locative agreement are semantically linked to the locative content of the phrase, unlike those bearing non-locative agreement. Thus in (40a), where the quantificational adjective *-iri -onse* – 'every' agrees with *chiseko* – 'door', the meaning is 'every door'; in (40b), where *iri -onse* bears locative agreement, it is interpreted as quantifying over places on the door, so the meaning is 'every part of the door, or 'all over the door':

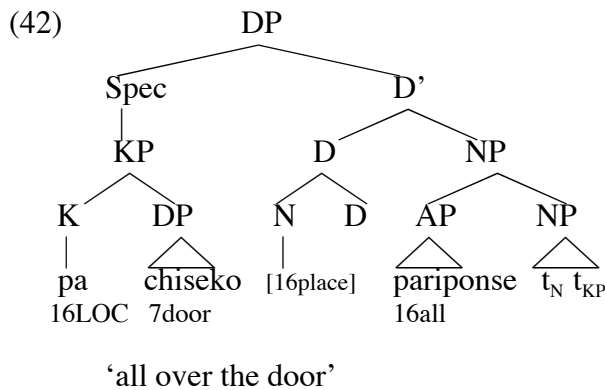
- (40) a. pa-li nchenche pa-chiseko chirichonse. [Chichewa]  
16-be 10fly 16-7door 7every  
'there are flies on every door'
- b. pa-li nchenche pa-chiseko pariponse.  
16-be 10fly 16-7door 16every  
'there are flies all over the door'



These facts argue that in addition to the overt noun stem within locative phrases there is also a sort of ‘place’ noun present, which takes modifiers of its own, and controls agreement on them. Bresnan and Mchombo (1995) analyze locative prefixes themselves as nouns, given that they seem to control agreement like nouns do, and based on evidence like conjoinability that they are independent words rather than true prefixes (see 41).

- (41) Ndi-na-jámbúl-a                      zithunzi mu ndí pa madengu.      [Chichewa]  
 1SG-SB-REC.PST-draw-IND 8picture 18 and 16 6basket  
 ‘I drew pictures in and on baskets.’

Carstens (1997) argues against the analysis of locative prefixes as nouns, since the prefixes violate the Two Mora Constraint to which nouns are subject, and fail to trigger ‘of’ insertion before their nominal complements (otherwise obligatory for complements of nouns). Functional categories are exempt from the Two Mora Constraint, so locative prefixes pattern as functional elements in this regard; ‘Of’-insertion does not apply between a Case-“assigning” category and its complement; thus the locative prefixes behave like Case-“assigners”. Carstens concludes that locative prefixes are gender-specific Case-markers, each selected by one of three silent ‘place’ nouns of distinct genders.<sup>23</sup> The silent locative noun adjoins to D as is standard in Bantu (see section 3). Word order indicates that the Case-marker and its complement (KP in 42) raise to the left edge of the locative DP, thus to its Spec. In this position the richly agreeing preposition licenses the empty noun by making its content recoverable (see 43):



- (43) An empty locative nominal is licensed by “rich” agreement (Carstens 1997)

## 2.5 Discussion

While Romance languages do not have locative genders, they do have well-documented sub-phrasal gaps which are disallowed in English, and which therefore seem to be connected with the availability of “rich” agreement on nearby categories:

(44) Me gusta el /la [e] de Juan [Spanish]  
 1sg.dat please the.masc/the.fem of  
 \*‘I like the of Juan’

(45) Q: Que libro quieres? [Spanish]  
 Which book want.2S  
 ‘Which book do you want?’

A: El grande [e].  
 the.masc big  
 \*‘The big’.

It seems likely that these gaps, like the null Bantu ‘place’ nouns, are available at least in part because agreement makes their content recoverable.<sup>24</sup> I follow Carstens (1997) in concluding that Bantu place nouns and these Romance gaps are subphrasal counterparts to *pro*; assuming that zero-morphemes are sub-word-level *pro*, the whole set of phenomena described in the section can receive a unified treatment; see the generalized licensing requirement in (46).

(46) The  $\phi$ -features of a silent nominal element must be identified.

Rich agreement (or, in the case of zero morphemes, an identifying feature of a “rich”gender system) licenses silent elements at all levels of structure. Bantu has more gender distinctions than Romance, and hence can identify unambiguously a greater number of null elements.

### 3. Word order and phrase structure in DP.

In this section I turn to DP-internal syntax, showing that Bantu and Romance DPs have a shared architecture and many common properties. Nouns move higher in Bantu than in Romance, however. In both languages, modifiers can be adjoined to the left or to the right of N’s base position, yielding variability in post-nominal modifier order.

#### 3.1 The noun-raising approach.

Bantu languages generally lack overt articles. In a DP containing a lexical possessor, agent, or theme argument in an ‘of’ phrase, standard word order is N-AP-ofP:<sup>25</sup> In contrast to lexical arguments, genitive pronouns are positioned between the noun and any APs:

(47) a. picha nzuri ya Halima [Swahili]  
 9picture 9nice 9of  
 ‘Halima’s nice picture’

Recent developments have led to a more complex picture of DP-syntax. It is well-known that languages in which modifiers precede the noun exhibit consistent ordering among the modifiers, while those with nouns left-peripheral in DP often have either the same order or its mirror image. These generalizations are expressed in Greenberg's Universal 20, reproduced below (cf. Greenberg 1963).

- (50) **Greenberg's Universal 20:** When any or all of the items (demonstrative, numeral, and descriptive adjective) precedes the noun, they are always found in that order. If they follow the noun the order is either the same or its exact opposite.

Universal 20 has proved to be too strong in that other post-nominal orders exist (see Cinque 2005 for extensive discussion). But the claimed inflexibility of pre-nominal modifiers has proved accurate. The contrast between pre- and post-nominal modifiers in this regard, and the various post-nominal orders, present a challenge addressed in much current research on noun phrase syntax.

Cinque (2000, 2005) argues that to obtain all and only the licit orders, it is necessary to assume that the pre-nominal order is universally base-generated:

- (51) *Universal Merge Order: Dem > Numeral > Adjective > N*

Cinque observes that if noun-raising accounted for all variation in word order, and if head-positions intervened between modifiers, we would expect only the following orders:

- (52) a. *Dem Numeral Adj N* (no movement)
- b. *N Dem Numeral Adj t* (complete N-raising)
- c. *Dem N Numeral Adj t* (N raises 2 levels)
- d. *Dem Numeral N Adj t* (N raises 1 level)

This leaves unaccounted for the common mirror-image order, and 9 additional possibilities that are attested (cf. Cinque 2005). Cinque accordingly adopts a different approach, arguing that there is no head-movement at all in grammar. A complex array of leftwards NP-movements and roll-up movements produce the attested orders among nouns and their modifiers (Cinque 2005:321-324).

With this much established I turn to modifier order within Bantu DPs.

Rugemalira (2006) reports that Swahili allows the orders Noun-Demonstrative and Demonstrative-Noun, with free word order among the modifiers to the right of these items (53a-d adapted from Rugemalira 2006:12-13). The order Numeral-Adjective in (53a) and (53c) is consistent with (52), but the order Adjective-Numeral in (53b, d) is a mirror image order (though a partial one, since the demonstrative is not final).

- (53) a. wale watu wawili wazuri [Dem-N-Num-Adj]  
           2those 2person 2two 2good
- b. wale watu wazuri wawili [Dem-N-Adj-Num]  
      2those 2person 2good 2two
- c. watu wale wawili wazuri [N-Dem-Num-Adj]  
      2person 2those 2two 2good
- d. watu wale wazuri wawili [N-Dem-Adj-Num]  
      2person 2those 2good 2two
- ‘those two good people’

If we apply the derivations in Cinque (2005) to the Swahili orders in (53), here is how they work. The Merge order is ‘those two good people’. Two surface orders are produced by NP moving all the way up (see 54), or partially, ending just to the left of the numeral (55) (I bracket only moved constituents):

- (54) [<sub>NP</sub> people] those two good  $t_{NP}$  (= 53c; *complete NP-movement*)
- (55) those [<sub>NP</sub> people] two good  $t_{NP}$  (= 53a; *partial NP-movement*)

The orders in (53b) and (53d) are derived by a sequence of movements. First, the NP raises to the left of the adjective. It does not surface in this position, however; the projection containing NP, adjective and the NP’s trace (= XP in 56) subsequently moves leftwards across the numeral:

- (56) a. those two [<sub>NP</sub> people] good  $t_{NP}$  →
- b. those [<sub>XP</sub> people good  $t_{NP}$ ] two  $t_{XP}$  (=53b)

(53d) follows up on this derivation with one final step, moving the NP alone again, this time across the demonstrative:

- (57) [<sub>NP</sub> people] those [<sub>XP</sub>  $t_{NP}$  good  $t_{NP}$ ] two  $t_{XP}$  (=53d)

Cinque acknowledges that motivation for the movements he posits remains to be found.

### 3.3 An alternative

In what follows I will sketch out an alternative approach to these phenomena that is more in keeping with surface appearances. I adopt the hypotheses in (58):

- (58) a. Spec, head, complement are Merged, always in the same hierarchical

and left-to-right arrangement (cf. Kayne 1994).

- b. Modifiers are Adjoined, always in the same hierarchical arrangement whether on the left or right edge of the constituent they modify.<sup>27</sup>

*Universal Hierarchy (adapted from 51):*

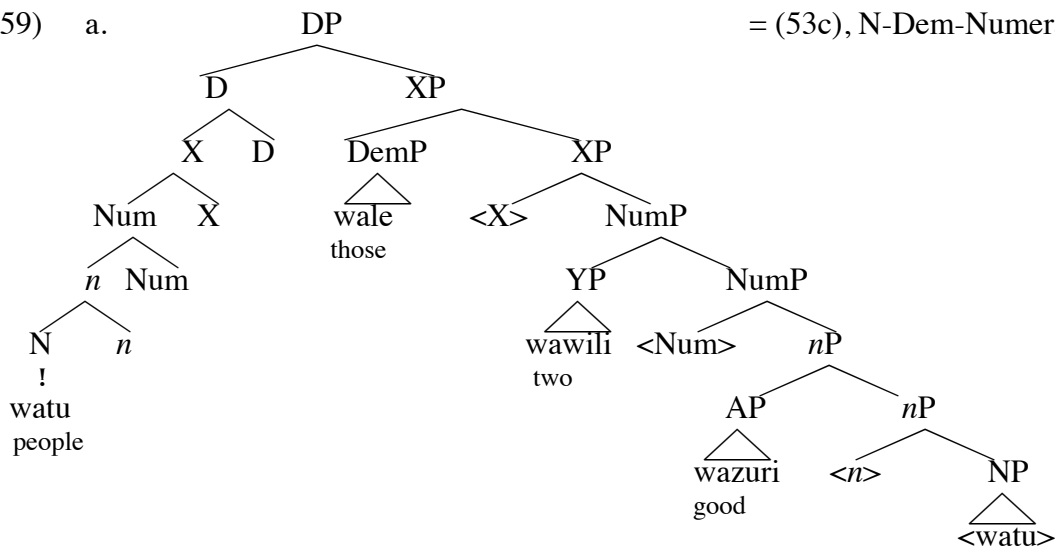
*Dem > Numeral > Adjective > N*

- c. Movement is always to the left.

(58) maintains a traditional distinction between the outputs of the Merge operation on the one hand, and an Adjunction operation adding modifiers, on the other; in the terminology of Chomsky (2000), these are the Set-Merge and Pair-Merge processes respectively. Under (58), syntax is weakly antisymmetric, in the sense of Takano (2003): Spec-head-complement material conforms to some version of the Linear Correspondence Axiom of Kayne (1994), mapping hierarchical relations into left-to-right order;<sup>28</sup> but adjuncts are distinct from specifiers, contra Cinque (1999, 2005), and can appear to the left or to the right edge of the constituent they modify. Hypothesis (58b) thus reflects the variability of modifier order in DPs, while positing hierarchical rigidity among modifiers. As for (58c), we may attribute it to movement's status as a variety of Set-Merge (cf. Takano op cit).

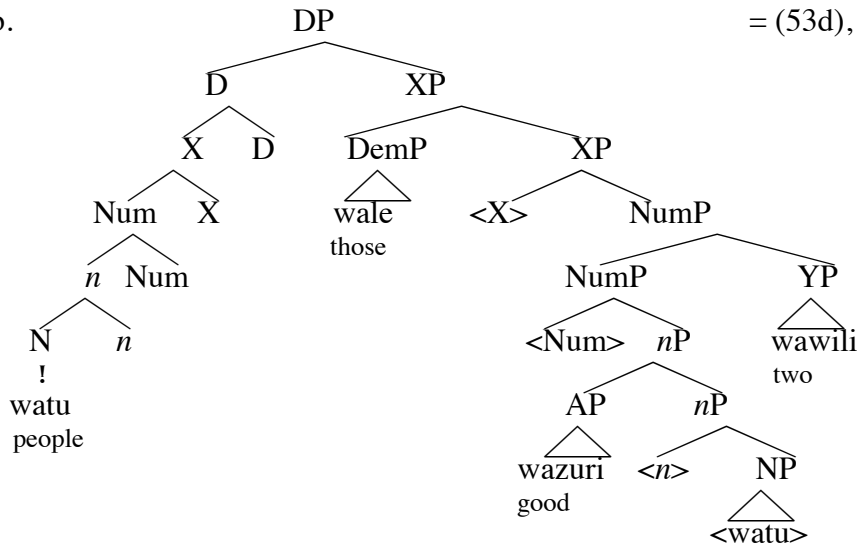
(59) sketches out representations for two of the acceptable Swahili word orders (53c,d). APs are adjuncts to *nP*, the nominal counterpart to *vP*. I leave open the identity of XP, the projection modified by demonstratives; for arguments that demonstratives are adjuncts below DP in Swahili, see Carstens (1991). I take numerals to be universally adjuncts to Num(ber)P; in Swahili, they have the flexibility to be adjoined either to the right or to the left (59a versus 59b).<sup>29</sup> Head-movement applies cyclically to move the noun to the left edge of DP.<sup>30</sup>

(59) a. = (53c), N-Dem-Numeral-Adj



b.

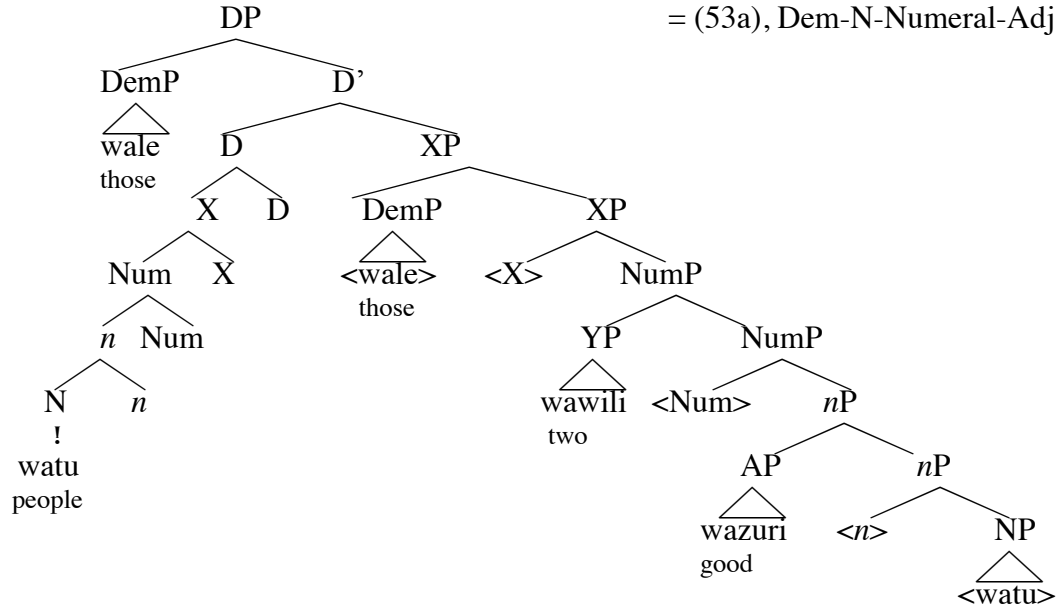
= (53d), N-Dem-Adj-Numeral



If a language permits demonstratives to adjoin on the right edge of XP, this will add two additional orders to the four described. Rugemalira 2006:14 describes these orders as alternates available in closely-related Safwa.

(53a,b) differ from (59a,b) respectively only in the appearance of the demonstrative in initial position. I follow Carstens (1991) in analyzing Bantu initial demonstratives as appearing in Spec, DP. Ashton (1944) states that Swahili pre-nominal demonstratives are semantically different from their post-nominal counterparts -- close in interpretation to definite determiners, picking out something afore-mentioned (Ashton 1944: 59 & 181-183). While it is in principle possible that they are slightly different lexical items with different Merge positions, there is a more intriguing interpretation that does not rely on a proliferation of homophonous lexical items. Just as definite DPs tend to move to high positions (cf. Diesing 1992), raising a definite demonstrative to Spec, DP plausibly correlates with a definite interpretation for it.

(60) = (53a), Dem-N-Numeral-Adj



### 3.4 Discussion

Symmetric base-generation of modifiers and some basic leftward movements account simply for the ordering possibilities among nouns and their modifiers in Swahili. These proposals fit into a constrained model of grammar: the ban on rightwards movement and the universal hierarchy governing adjunction proposed in (58b) restrict the possibilities to only those that are found in the languages of the world (see Cinque 2005: 319-20). Assuming also that movement respects Relativized Minimality (Rizzi 1990) or its minimalist successor Attract Closest, the universal hierarchy is consistently maintained.

Cinque describes fourteen attested orderings of the elements Noun, Demonstrative, Adjective, and Numeral. Assuming that head positions intervene between each modifier, where nouns might surface in different languages, twelve word orders are generated by the combination of symmetric modifier adjunction possibilities and noun-movement. Only two orderings in the languages of the world have no obvious account in terms of variable adjunction direction and head-movement: [Adj-N-Dem-Numeral] and [N-Adj-Dem-Numeral]. These would seem best accounted for by raising the NP containing Adjective and Noun as Cinque proposes, in a kind of predicate-fronting comparable to the clause-level operation of VP-raising.

### 3.5 Romance DPs

Adjectives in Romance are often found post-nominally, and the same issues arise in connection with their relative ordering. Laenzlinger (2005) reports that postnominal APs in French exhibit either the sequence in (61a) or mirror image order (62a). Laenzlinger, following Cinque's approach, concludes that French sometimes moves the smallest NP ([<sub>NP</sub> chaise] in 60b), but can also do roll-up movement as shown in (62) (moved constituents bracketed).



(61) a. une chaise ronde<sub>1</sub> rouge<sub>2</sub> [French]  
a chair round red

b. une \_\_\_\_ ronde rouge [chaise]  
↑  
└──────────┘

(62) a. une voiture rouge<sub>2</sub> splendide<sub>1</sub> vs. a beautiful<sub>1</sub> red<sub>2</sub> car

b. une splendide \_\_\_\_ rouge [voiture] [French]  
↑  
└──────────┘

c. une \_\_\_\_ splendide [voiture rouge t]  
↑  
└──────────┘

As in the case of Bantu, an alternative analysis exists; it seems in principle possible that *splendide*, though hierarchically superior to *rouge*, can be Adjoined to the right.

Future research will no doubt uncover compelling reasons to reject one of these approaches to word order variation in grammar. This review of the theoretical issues and the Bantu and Romance facts illustrates how closely parallel are the structures and processes of these genetically unrelated groups of languages. We can be confident, I think, that a single theoretical model will account for them all.

#### 4. Agreement in DP.

The final topic of this comparative paper is the phenomenon of agreement or concord among the elements within DP. In both Bantu and Romance languages, modifiers and arguments in DP inflect for the gender and number features of the head noun.

(63) a. kikombe changu cheupe [Swahili]  
7cup 7my 7white  
'my white cup'

b. vikombe vyangu vyeupe  
8cup 8my 8white  
'my white cups'

(64) a. mtoto huyu mdogo  
1child 1this 1small  
'this small child'

b. watoto hawa wadogo  
2child 2this 2small  
'these small children'

- (65) a. el niño pequeño [Spanish]  
the.masc child.masc small.masc  
'the small boy'
- b. la niña pequeña  
the.fem child.fem small.fem  
'the small girl'

In the framework of Chomsky (1986) and (1992), agreement was analyzed as a Spec-head relation (see also Koopman 1992). Thus subject agreement occurs because the features of a DP in Spec, TP are shared with the head T(ense).

- (66) a.  $[[_{\text{Spec TP}} \text{a man}] \text{T+be in the room}]$   
b.  $[[_{\text{Spec TP}} \text{a man}] \text{T+be+3SAgr in the room}]$   
c. A man is in the room

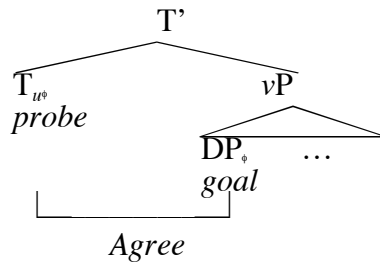
This approach did not readily extend to concord in DP. Based on the Bantu pattern of concord Carstens (1991) accordingly argued that agreement could be licensed either through Spec-head feature-sharing or through feature-sharing in the government relation, the latter yielding concord. Such a disjunctive approach is theoretically undesirable, but seemed unavoidable.

But in Minimalist theory, agreement is licensed in the Agree relation (Chomsky 2000; 2001, which is based on closest c-command. The agreement features (the *probe*), must c-command the category they agree with (the *goal*):

- (67) a. Agree ( $\alpha, \beta$ ) iff  $\alpha$  c-commands  $\beta$ ;  $\alpha, \beta$  have matching features; there is no  $\gamma$  with matching features which is closer to  $\alpha$  than  $\beta$ .
- b.
- 

The canonical case again is subject agreement (SA). T(ense) enters the syntax bearing the unvalued, uninterpretable  $\phi$ -features that underlie SA. The closest potential goal for these is the subject in its base position, Spec,  $\nu$ P. The subject therefore provides T's  $\phi$ -features with values for person, number, and, in Bantu, gender.

(68)



English expletive constructions illustrate that closest c-command underlies agreement, since subject agreement can reflect the features of a DP that never raises to Spec, TP, since an expletive is there (see 69). Alternatively, the agreed-with category can move to Spec, TP (see 70) to satisfy the requirement that TP have something in that position (= the EPP). But agreement is available whether subject raising applies or not.

(69) a. There  $T_{t\phi}$  be [two men] in the room  
Agree

b. There  $T_{3PL}$ +be [two men] in the room

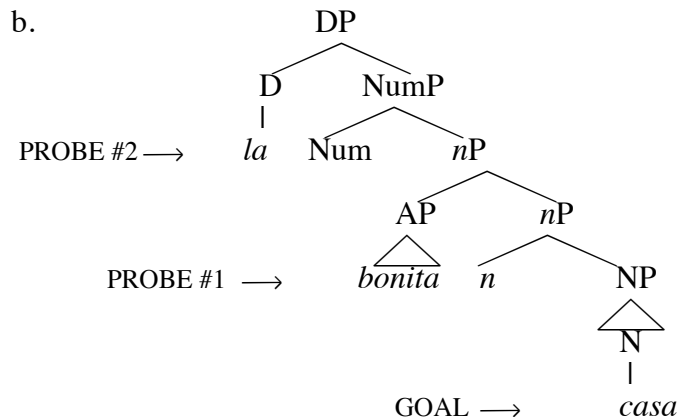
c. There are two men in the room.

(70) a. Two men are in the room

b. [Two men]  $T_{3PL}$ +be  $t_{two\ men}$  in the room

Returning to the topic of concord, Carstens (2001) shows that Agree accounts for it straightforwardly. The agreeing categories in DP always c-command the head noun in its Merge position, prior to noun-raising; hence Agree predicts that their features can licitly be valued.

(71) a. la casa bonita [Spanish]  
The(f) house(f) pretty(f)  
'the pretty house'



Just one theoretical modification is needed for Agree to account for concord in Universal Grammar. Chomsky (2000, 2001) proposes that a goal must have an undeleted Case feature, and that every Agree relation that values  $\phi$ -features also deletes Case. If this were true, multiple agreement with a single category would be impossible. I follow Carstens (2001, 2003, 2005) in concluding that only a head lexically specified with the Case-“assigning” property deletes Case in the Agree relation. Under this view, concord is an entirely unexceptional instance of agreement in grammar.

## 5. Conclusion.

Resemblances between Bantu and Romance languages have always been apparent in their gender systems and extensive agreement phenomena. In this paper I have looked closely at several aspects of DP syntax in languages from the two groups, including their nominal gender systems; word order and phrase structure in DP; DP-internal concord; and the ability of such concord to license sub-phrasal nominal gaps. The comparison and the results are summarized in (72):

| (72) <i>Phenomena</i>      | <i>Romance</i>                                                | <i>Bantu</i>                                            |
|----------------------------|---------------------------------------------------------------|---------------------------------------------------------|
| <i>Genders:</i>            | 2-3                                                           | Numerous (in Swahili, 6 plus locative, gerundive, etc.) |
| <i>Nominal Morphology:</i> | Suffixal, often concatenating gender and number markers       | Prefixal and fusional                                   |
| <i>DP structure:</i>       | Identical for the two families                                |                                                         |
| <i>N-raising:</i>          | To DP's middle field                                          | To D                                                    |
| <i>Concord:</i>            | For both families, in gender & number, via the Agree relation |                                                         |
| <i>Empty nouns:</i>        | Licensed by concord in both families                          |                                                         |

Thus close investigation confirms that Bantu and Romance families are indeed highly similar in many respects. This in turn confirms the validity and utility of Universal Grammar as a hypothesis and a set of analytical tools.

## References

- Ashton, E.O. 1944. *Swahili Grammar Including Intonation*. London: Longman Group Ltd.
- Bresnan, J. & S. Mchombo. 1995. The lexical integrity hypothesis: evidence from Bantu. *Natural Language and Linguistic Theory* 13: 181-254.
- Carstens, V. 1991. The morphology and syntax of determiner phrases in Kiswahili. UCLA PhD thesis.

- Carstens, V. 1993. On nominal morphology and DP structure. In *Theoretical Aspects of Bantu Grammar*, S. Mchombo (ed.). Stanford: Stanford University Press.
- Carstens, V. 1997. Empty nouns in Bantu locatives. *The Linguistic Review* 14: 361-410.
- Carstens, V. 2001. Multiple agreement and Case deletion: Against  $\phi$ -(in)completeness. *Syntax* 3:147-163.
- Carstens, V. 2003. Rethinking complementizer agreement: Agree with a Case-checked goal. *Linguistic Inquiry* 34.3: 393-412.
- Carstens, V. 2005. Agree and EPP in Bantu. *Natural Language and Linguistic Theory* 23: 219-279.
- Chomsky, N. 1986. *Barriers*. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1992. Some notes on economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*, R. Friedin (ed), 417-454. Cambridge, Mass.: MIT Press.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, Mass: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: the framework. In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, H. Lasnik, R. Martin, D. Michaels, and J. Uriagereka (eds.), 89-156. Cambridge, Mass: MIT Press.
- Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A Life in Language*, M. Kenstowicz (ed.), 1-52. Cambridge: MIT Press.
- Cinque, G. 1995. *Italian Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Cinque, G. 1999. *Adverbs and Functional Heads*. New York: Oxford University Press.
- Cinque, G. 2000. On Greenberg's Universal 20 and the Semitic DP. In *University of Venice Working Papers in Linguistics*, L. Bruge (ed), 10.2:45-61.
- Cinque, G. 2005. On Deriving Greenberg's Universal 20. *LI* 36.3: 315-332.
- Contini-Morava, E. 2000. Noun class and number in Swahili. In *Between Grammar and Lexicon*, E. Contini-Morava and Y. Tobi, (eds), 3-30. Amsterdam: John Benjamins.
- Contreras, H. 1989. On Spanish empty N' and N. In *Studies in Romance Linguistics*, C. Kirschner and J. DeCesaris (eds.), 83-95. Amsterdam: John Benjamins.
- Corbett, G. 1991. *Gender*. Cambridge: Cambridge University Press.

- Diesing, M. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.
- Di Sciullo, A.M. 2005. *Asymmetry in Morphology*. Cambridge, Mass.: MIT Press.
- Ferrari, F. 2005. A Syntactic Analysis of the Nominal System of Italian and Luganda – how nouns can be formed in the syntax. NYU PhD dissertation.
- Greenberg, J. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, J. Greenberg (ed), 73-113. Cambridge, Mass.: MIT Press.
- Halle, M. and A. Marantz. 1993. Distributed morphology and the pieces of inflection. In *The View from Building 20*, K. Hale and S.J. Keyser (eds), 111-176. Cambridge: MIT Press.
- Harris, J. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22.1: 27-62.
- Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Koopman, H. 1992. On the absence of Case-chains in Bambara. *Natural Language and Linguistic Theory* 10: 555-594.
- Laenzlinger, 2005. French adjective ordering: perspectives on DP-internal movement types. *Lingua* 115: 645-689.
- Lubowicz, A. 2002. Derived environment effects in Optimality theory. *Lingua* 112: 243-280.
- McCarthy, J. 2003. Comparative markedness. *Theoretical Linguistics* 29: 1-51.
- Meinhoff, C. 1899. Grundriss einer Lautlehre der Bantusprachen. Berlin: Deutsche Morgenländische Gesellschaft, Abhandlungen für die Kunde des Morgenlandes.
- Mufwene, S. 1980. Bantu Class prefixes: Inflectional or derivational? *Papers from the 16<sup>th</sup> Regional Meeting*. Chicago: CLS.
- Mutashansky, O. 2006. Heads-movement in linguistic theory. *Linguistic Inquiry* 37.1:69-109.
- Myers, S. 1987. *Tone and the Structure of Words in Shona*. U. Mass Amherst PhD dissertation.
- Njogu, K. 1994. The minimal size condition in KiSwahili. Paper presented at the 25<sup>th</sup> Annual Conference on African Linguistics, Rutgers University, New Brunswick.

Park, Jae-Ik. 1997. Disyllabic requirement in Swahili morphology. U. Penn Working Papers in Linguistics, Vol. 4.2: 245-259.

Pesetsky, David. 1990. Experiencer Predicates and Universal Alignment Principles. Ms., MIT.

Ritter, E. 1988. A head-movement approach to construct state noun phrases. *Linguistics* 26: 909-929.

Ritter, E. 1992. Cross-linguistic evidence for Number Phrase. *The Canadian Journal of Linguistics* 37.2: 197-218.

Rizzi, L. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.

Roeper, T. 1999. Leftward movement in morphology. In *MIT Working Papers in Linguistics* 34: 35-66.

Rugemalira, J. 2006. The structure of the Bantu noun phrase. Ms., Dar es Salaam.

Selkirk, E. 1982. *The Syntax of Words*. Cambridge, MA: MIT Press.

Sproat, R. 1985. On Deriving the Lexicon. MIT PhD. dissertation.

Takano, Y. 2003. How antisymmetric is syntax? *Linguistic Inquiry* 34.3: 516-526.

Valois, D. 1991. The Internal Syntax of DP. UCLA PhD. Dissertation.

---

\* My thanks to Josephat Rugemalira for helpful discussion of the Swahili data in section 3 of this paper.

<sup>1</sup>Harris (1991) argues convincingly that Spanish word markers should not be viewed as gender morphology; rather, their distribution is determined by redundancy rules in which gender plays a limited role. I abstract away from this and gloss the markers *-a* and *-o* as F(eminine) and M(asculine) respectively, as a convenience, while assuming that gender is a lexical property of nouns. Also, I will not discuss the presence of neuter gender in some Romance languages.

<sup>2</sup> Bantu languages differ as to which and how many of the proto-Bantu classes are preserved. There are also language-particular deviations from the pairings 1/ 2, 3/ 4, etc.: Swahili nouns in Class 11 take their plurals in Class 10, proto-Bantu Class 12 being absent; a reviewer points out that in Sesotho some Class 1 stems take plurals in Class 6; some Class 5 ones take plurals in both Class 6 and 10; and so forth.

<sup>3</sup> Italian plurals are rather like Bantu Class-marked nouns in that the morphology of an Italian plural noun correlates fairly predictably with the noun's grammatical gender, and there is no separate word marker for gender on plurals: *casa/case* – house/s(f), *pizza/pizze* – pizza/s(f); versus *ragazzo/ragazzi* – boy/s(m), *tempo/tempi* – time/s(m). It is nonetheless standard to analyze interpretable number as a feature distinct from gender, the latter a lexical property of Italian nouns. My analysis extends this reasoning to Bantu.

<sup>4</sup> Though I adopt this analysis, my glosses will continue to follow the traditional numbering system for consistency with Bantuist practice.

---

<sup>5</sup>Though I use the term *stem* to designate the form that Noun Class prefixes attach to, I assume that gender is a lexical property of noun roots. This follows from the fact that the gender is largely arbitrary; it must be part of the lexical listing, since it cannot be predicted (see discussion of 18).

<sup>6</sup> An alternative is to suppose that gender and number are two distinct morphemes underlyingly; but in Bantu, gender and number undergo the operation *fusion* (Halle & Marantz 1993). Since I am assuming gender is strictly a lexical property of nouns I leave this possibility aside.

<sup>7</sup> In Gender D, all forms of agreement are homophonous with the prefixes on nouns (save for some phonologically-conditioned allomorphy). This is not the case in all genders; for example, in Gender A singular, subject agreement is *a-* rather than *m-*, and the Class morphology for demonstratives is a third form, *yu-*. Category-specific Spell-Out rules like these are therefore needed.

<sup>8</sup> In Class 3 (Gender B Singular), though the Noun Class prefix is *m-*, the subject agreement prefix is *u-*. See note 7.

<sup>9</sup> My thanks to an anonymous reviewer for pointing these examples out.

<sup>10</sup> *Ji*-epenthesis is also employed in two other contexts. First is the formation of super-diminutives and super-augmentatives, a fact that Carstens (1991) relates to the Zero-Morpheme licensing requirement (see 28), since a doubly zero-affixed root needs a means to identify the outer zero morpheme and distinguish it from the inner one. Second there is a constraint prohibiting vowel-initial nouns and adjectives that can be satisfied by *ji*-epenthesis. See Carstens (1991:219-238) for details.

<sup>11</sup> In the Optimality framework there are no derivations; hence derived environment effects are explained without reference to derivations. See Lubowicz (2002), McCarthy (2003) and references cited there. I use the traditional term as an expository convenience.

<sup>12</sup> An anonymous reviewer points out that diminutives are formed with overt suffixes in some Bantu languages.

<sup>13</sup> It has been argued that words have underlying (Spec)-head-complement structure and order (see Kayne 1994, Roeper 1999, Di Sciullo 2005 for various implementations of this idea). Applied to the case under consideration, this would mean that the diminutive affix is Merged to the left of the stem, and the stem reaches its surface position by leftwards movement. This interesting possibility has no obvious bearing on the issues I address here (the nature of Class, and its non-involvement in derivation) so I will leave it aside.

<sup>14</sup>In the reconstructed proto-Bantu, diminutives belonged to special Classes, usually 12/13, and this system persists in many of the modern languages including Kikuyu. Note that Kikuyu diminutive formation is also subject to the 2 Mora minimum word requirement; a monomoraic stem retains the prefix of its lexical class as an inner prefix, while most polymoraic stems do not. Thus the zero-morpheme analysis is motivated for a language with Class 12 diminutives.

|     |    |                                       |   |                               |
|-----|----|---------------------------------------|---|-------------------------------|
| (i) | a. | ka-iriitu<br>12-girl<br>'little girl' | < | mu-iriitu<br>1-girl<br>'girl' |
|     | b. | ka-mu-ti<br>12-3-tree<br>'stick'      | < | mu-ti<br>3-tree<br>'tree'     |

There is evidence that minimum word requirements do not apply cyclically in all languages (see discussion of Navajo in Carstens 1991: 239-244). It would therefore not be surprising to find Bantu languages in which zero-affixation of the diminutive morpheme did not lead to epenthesis in the stem; this would not indicate that the zero-affixation approach is incorrect for such languages.



<sup>15</sup> There are ways of getting around this problem. Many Class 5/6 nouns form augmentatives in Classes 3 and 4; the epenthetic “inner prefix” morpheme *ji-* salvages otherwise impossible diminutives and some of the blocked augmentatives. Thus while *kitabu* – ‘book’ has no diminutive reading, *kijitabu* – ‘little book’ is fine. Carstens (1991:223-238) broadens the *ji*-epenthesis rule so it is triggered by violations of nominal well-formedness conditions generally, including the zero-morpheme licensing principle.

<sup>16</sup> One might alternatively conceive of the derivation of Swahili diminutives and augmentatives as a rule or process adding the relevant gender feature, rather than as a morpheme. But since, as noted in the text, Genders C and D do not have diminutive and augmentative meanings, their addition to nouns cannot suffice to indicate the derived noun’s altered semantics; there must additionally be a specification as diminutive or augmentative introduced:

- (ii) a. Diminutive formation:  $N_{\text{gender}} \rightarrow N\text{-DIMIN}_{\text{Gender D}}$   
 b. Augmentative formation:  $N_{\text{gender}} \rightarrow N\text{-AUGMENT}_{\text{Gender C}}$

I take these specifications to be indistinguishable from zero-morphemes; thus the analyses are notational variants.

<sup>17</sup> An appealing alternative idea is that the Noun Class prefixes themselves identify the zero-morphemes. But this approach falls short of explaining the facts in (iii); only a gender feature distinct from that of the stem suffices.

- (iii) a. *ok* *ki-*[*tabu*<sub>Gender D</sub>]  
           7- book  
           ‘book’  
 b. \* *ki-*[*tabu*<sub>Gender D</sub>]  $\emptyset$ <sub>Gender D</sub>  
           7- book DIMIN  
           ‘little book’

<sup>18</sup> The proposed percolation from non-heads of complex words has no correlate in Minimalist syntax, as an anonymous reviewer points out. Since the final vowel of the noun in Romance correlates fairly regularly with gender, an alternative account comes to mind, namely that the gender feature of Romance diminutives is added in the outermost layer, in the suffix, and this is why the diminutive affix itself has no impact on the gender of the derived form. But the suffix and the nominal root must match for gender specification, a fact that leads to the same conclusion I have argued for – when the diminutive is affixed to a nominal root, the gender of the latter determines the gender of the derived form. Hence the diminutive affix itself must lack a gender feature, and the stem’s gender percolate.

<sup>19</sup> An anonymous reviewer points out that gender variation with a single root in some Romance languages can be related to size dimension (i.e. Spanish *barca-barco* “small boat-boat”). Yet, both forms can also take diminutive-augmentative suffixes (*barquita/barcaza; barquito/barcote*). The suffixes are represented as in (30e). An account of connections between gender and semantics is sketched out in (38).

<sup>20</sup> Ferrari (2005) argues that Romance has derivational uses of gender like those she posits for Bantu. I consider these briefly below.

<sup>21</sup> Ferrari follows Contini-Morava (2000) in using the term “animate” to designate “entities with vitality”, that is, having to do with life or its sustenance; hence *granito* – ‘stone’ is inanimate but the grated-ice drink *granita* is animate because it is a food and food is life-sustaining. I find the over-arching semantic classifications unconvincing and it is not clear to me, if these components of nominal semantics reside in the genders, how it happens that a food item like *pomodoro* – ‘tomato’ is masculine. Thanks to an anonymous reviewer for this example.

---

<sup>22</sup>This rule is much like Harris's (1991) Human Cloning Rule, which operates on any noun with the semantic specification 'human' and no specification for sex. Human Cloning turns such nouns into pairs of nouns specified 'male' and 'female'. Then a redundancy rule 'female' → f/[\_\_human] maps female sex to feminine grammatical gender.

<sup>23</sup> An anonymous reviewer points out that Case-markers might reasonably be viewed as having only uninterpretable content; thus by the logic of section 2.1, KPs could be argued to lack syntactic legitimacy, like AgrPs. I accordingly assume that they have some prepositional semantic content (like 'in' or 'at') compatible with the meanings of the null nouns they combine with.

<sup>24</sup> An anonymous reviewer points out that agreement does not suffice in every case, since \**Me gusta el* – 'I like the' is ill-formed. Contreras (1989) attributes this to a clitic property of the determiner.

<sup>25</sup> Judgements vary on DPs containing two lexical arguments in 'of'Ps. Though several speakers I have consulted accept them, others find them degraded (cf. Rugemalira p.c.). Pursuit of this topic lies outside this paper's scope.

<sup>26</sup> The facts are more complex than this brief sketch suggests. In French and Spanish, articles and genitive pronouns cannot co-occur. In Italian, many adjectives are pre-nominal. While these points of variation are significant, space considerations preclude exploration here. See Cinque (1995), Valois (1991) for in-depth discussion of the Italian and French facts respectively.

<sup>27</sup> This would mean that the LCA of Kayne (1994) does not apply to adjuncts. I comment below.

<sup>28</sup> There have been several alternatives to the original LCA proposed in the syntactic literature; I will not explore these here. Representations that follow are not intended to conform to the LCA in all technical details, but assume that Spec-head-complement order of Merged material is universal.

<sup>29</sup> In their concordial prefixes, numerals pattern with adjectives. I leave their category an open question.

<sup>30</sup> See Matushansky (2006) for an analysis of head-movement as raising to a Spec of the root, followed by a morphological merger process (m-merger).