Noun-Modifier Order in Africa

Matthew S. Dryer
University at Buffalo

Many VO languages of Africa place all or most modifiers after the noun, as exemplified by Fongbe in (1).

(1) Fongbe (Lefebvre and Brousseau 2002: 54, 42)

a. àvɔ̀ ɔ̀vɔ̀ wè ɔ̀ lè
   loincloth red two DEF PLUR
   N Adj Num Def
   ‘the two red loincloths’

b. sùnù élɔ̀
   man DEM
   ‘this/that man’

Many Africanists, when confronted with languages like Fongbe, have assumed that this is because these languages are VO languages. However, as shown by Dryer (1992), it is not in general the case that modifiers of the noun follow the noun in VO languages, contrary to widely held beliefs. While adjectives (like ɔ̀vɔ̀ ‘red’ in 1a) more often follow the noun in VO languages, they do so just as often in OV languages. Numerals (like wè ‘two’ in 1a) in fact precede the noun more often in VO languages, and, outside of Africa, do so considerably more often than they do in OV languages. Definite markers (like ɔ̀ in 1a) precede the noun significantly more often in VO languages than they do in OV languages. And demonstratives (like élɔ̀ in 1b) precede the noun more often in both VO and OV languages (though they do so slightly more often in OV languages). In short, the fact that the modifiers follow the noun in Fongbe is not at all expected typologically.

The fact that modifiers follow the noun in Fongbe is not unexpected, however. Although not expected typologically, it is expected geographically. The primary thesis of this paper is that languages in Africa exhibit a greater tendency to place modifiers after nouns than languages in other parts of the world. Although this claim has been made before (Heine and Leyew 2007), this paper documents this tendency thoroughly, based on a large database containing data from over 1600 languages (Dryer 1992, Dryer 2005a, 2008, inter alia) and by comparing Africa explicitly with other areas in the world. This paper also shows that this tendency is exhibited by both VO and OV languages in Africa. Finally, it examines the distribution of this tendency within Africa and argues that it

---

1 Part of the research for this paper was made possible by Social Sciences and Humanities Research Council of Canada Grants 410-810949, 410-830354, and 410-850540 and by National Science Foundation Research Grant BNS-9011190.
clearly exists throughout Niger-Congo, Nilo-Saharan and Chadic and to a lesser extent among other Afroasiatic languages and among Khoisan languages (although among Khoisan languages it is primarily Khoe languages that do not conform).

1. Comparing Africa to the rest of the world

1. Order of adjective and noun

The data in this paper is presented in two ways, in maps and in bar graphs showing the frequency of different language types in different areas of the world (and, later in the paper, in different families in Africa). Consider first the order of adjective and noun. Map 1 shows the distribution of the two orders of adjective and noun throughout the world.

On Map 1 and on all subsequent maps, the black dots represent languages with postnominal modifiers while the white dots represent languages with prenominal modifiers. Map 1 shows the two orders of adjective and noun: the black dots show NAdj languages while the white dots show AdjN languages. It is clear from Map 1 that the languages of Africa are predominantly NAdj.

The second way I will present data is quantitative. In Table 1 is given data on the two orders of adjective and noun.

---

2 Versions of the maps shown in this paper are also available from the online version of the World Atlas of Language Structures (Haspelmath et al (eds.)) at http://www.wals.info/, where zooming in on Africa allows the user to identify particular languages. Map 1 does not include languages with both orders of adjective and noun where neither order is clearly dominant. The same applies analogously to all maps and data in this paper; only languages where one of two orders is dominant are considered.
The languages of the world are divided into six large continental areas: Africa, Eurasia (which excludes Sino-Tibetan languages and other languages of southeast Asia), Southeast Asia & Oceania, Australia - New Guinea, North America, and South America. The numbers represent, not numbers of languages, but numbers of genera, language groups comparable to the standard subfamilies of Indo-European, such as Germanic and Celtic (Dryer 1989, 1992, 2005b). Thus, the 15 on the first line under ‘Africa’ means that my database contains 15 genera in Africa that contain AdjN languages, while the 63 below it means that my database contains 63 genera in Africa that contain NAdj languages. The third line gives the proportion of genera in each area that contain NAdj languages as opposed to AdjN languages, i.e., the number on the second line as a proportion of the sum of the number on the first line and the number on the second line, the proportion of genera in each area that contain NAdj languages as opposed to AdjN languages. In the case of Africa, this figure is 81%. Throughout this paper, I will be giving these proportions in the form of bar graphs, as in Fig. 1, which represents the figures on the third line in Table 1, that is, the proportions in each of the six areas, as well as the average of these six proportions, which provides the best estimate of the likelihood of a language being of the given type. For example the average of 62% in Fig. 1 means that the overall likelihood of a language placing adjectives after the noun is 62%, reflecting an overall preference for NAdj order over AdjN order.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Africa</th>
<th>Euras</th>
<th>SEA&amp;Oc</th>
<th>Aus-NG</th>
<th>N.Amer</th>
<th>S.Amer</th>
<th>TOTAL</th>
<th>#Lgs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdjN</td>
<td>15</td>
<td>36</td>
<td>14</td>
<td>21</td>
<td>33</td>
<td>19</td>
<td>138</td>
<td>373</td>
</tr>
<tr>
<td>NAdj</td>
<td>63</td>
<td>7</td>
<td>47</td>
<td>79</td>
<td>30</td>
<td>42</td>
<td>268</td>
<td>870</td>
</tr>
<tr>
<td>Prop. NAdj</td>
<td>81%</td>
<td>16%</td>
<td>77%</td>
<td>79%</td>
<td>48%</td>
<td>69%</td>
<td>Avg=62%</td>
<td></td>
</tr>
</tbody>
</table>

3 A list of genera is also available online under each family at http://wals.info/lan.
4 Note that in a few cases, a genus will contain languages of both types, in which case it is counted on both lines. For that reason, the total of the figure on the first line and the figure on the second line may be greater than the total number of genera in the area and one cannot literally say that 81% of the genera in Africa contain NAdj languages. Strictly speaking, the number of genera containing NAdj languages in Africa is 81% of the sum of the number of genera containing AdjN languages and the number of genera containing NAdj languages. I will simply refer to this as the proportion of genera, with the understanding that it should not be taken literally, but should be understood in the technical sense explained in this footnote.
Fig. 1 shows that the proportion of genera containing NAdj languages is higher in Africa (81%) than in the other five areas, though only marginally higher than in two other areas, but considerably higher than the world average of 62%.

It is worth showing that this higher frequency of NAdj order in Africa is true for both VO and OV languages. Consider first VO languages. Fig. 2 shows the proportion of genera that contain NAdj languages among VO languages, and here the proportion for Africa is clearly higher than in any other area, at 86%, with the next highest proportion being 75% and the world average 59%. Fig. 3 shows that the proportion of NAdj order among OV languages and shows that Africa is among three areas with approximately the same proportion that is highest, all around 80%, well above the world average of 65%. Note that as I have shown previously (Dryer 1988, 1992), there is no tendency for OV languages to be AdjN; in every area other than Eurasia, OV languages are more often NAdj, and the average proportion for NAdj is actually slightly higher for OV than it is for VO (65% vs. 59%).

1.2 Order of demonstrative and noun

Map 2 shows the distribution of the two orders of demonstrative and noun among the languages of the world. Again we see that postnominal order, represented by the black dots, is the dominant type in Africa.
Fig. 4 shows that the proportion of genera that are NDem is higher in Africa than in the other five areas, at 74%, which is twice the world average of 37%.

Furthermore, this higher frequency of NDem order in Africa is again found among both VO and OV languages. Fig. 5 shows that the proportion of genera in Africa containing NDem order among VO languages is 83%, higher than the other five areas, and again much higher than the world average of 43%.
Fig. 5
Proportions of genera containing NDem languages among VO languages

Fig. 6
Proportions of genera containing NDem languages among OV languages

Fig. 6 shows that the proportion for Africa of NDem among OV languages, though lower than among VO languages at 57%, is still the highest among the six areas and much higher than the world average of 30%.

1.3 Order of numeral and noun

Map 3 shows the two orders of numeral and noun. Again, postnominal order clearly dominates in Africa.

Fig. 7 shows that the proportion is again highest in Africa (80%), though only marginally higher than Australia-New Guinea, but still nearly twice the world average of 43%.
Fig. 7
Proportions of genera containing NNum languages

Map 4 shows the distribution of the two orders of numeral and noun among VO languages. Not only does it show that NNum order is dominant in Africa, especially sub-Saharan Africa, but this order is relatively infrequent elsewhere in the world. In fact, my database contains only five VO&NumN languages in sub-Saharan languages while it contains only one language of the opposite VO&NNum type in all of the Americas and Eurasia combined (in the technical sense of Eurasia assumed here that excludes southeast Asia).

Fig. 8 shows that the proportion of genera containing NNum languages is again highest in Africa at 86%, much higher than the world average of 31%.
Fig. 9 shows analogous data for OV languages. This is one of the few cases in this paper where the proportion is noticeably lower in Africa than it is in some other area, since the proportion for Africa here is 73%, compared to 87% for Southeast Asia & Oceania and 85% for Australia-New Guinea. Nevertheless, the proportion for Africa is still higher than the world average of 55%.

1.4 Order of genitive and noun

The next noun modifier examined, that of genitives, is one whose order with respect to the noun clearly does correlate with the order of object and verb, in that OV languages tend to be GenN, while VO languages tend to be NGen. Map 5 shows clearly how GenN order is far more common than NGen order among OV languages.

However, if one looks at where the exceptional OV&NGen languages occur on Map 5, one finds that there are more of them in Africa than in most other areas of the world. And this become clearer in Fig. 10, where the proportion for Africa (31%) is far higher than for any other area, the next highest being 10%.
While OV languages tend to be GenN and while verb-initial languages tend to be NGen, SVO languages are split between the two orders of genitive and noun, as shown by Dryer (1992), so the order of genitive and noun is not predictable for SVO languages and it is thus worth specifically examining the distribution the two orders among SVO languages, as shown on Map 6.

Inspection of Map 6 shows that the postnominal order, NGen, is the more common order in Africa. And Fig. 11 shows that the proportion for NGen among SVO languages is higher in Africa (81%) than in the other five areas and much higher than the world average of 51%.
1.5 Order of relative clause and noun

Map 7 shows the two orders of noun and relative clause in OV languages. The order of relative clause and noun does correlate with the order of object and verb, but this is due to the fact VO languages are almost exclusively NRel (Dryer 1992, 2005c) while both orders are about equally common among OV languages.

But again, Map 7 shows that most of the OV languages of Africa are NRel. Fig. 12 shows that the proportion of NRel among OV languages is much higher for Africa (78%) than the world average (53%), though there is one area, North America, where the proportion is higher (84%).
1.6 Order of definite marker and noun

The last modifier I will examine is definite markers (or articles). As shown in Dryer (1992) for articles in general, the order of article and noun correlates with the order of object and verb in that article-noun order is more common among VO languages than it is in OV languages. Although the difference is not a great one, the difference is statistically significant. Map 8 shows the distribution of the two orders of definite marker and noun among VO languages. This map shows that not only is NDef clearly the most common order in Africa, but Africa is the area where this area is most common.

Map 8
Order of Definite Marker and Noun

Fig. 13 shows that the proportion of NDef among VO languages is highest in Africa (73%) and much higher than the world average (40%).

---

5 In Dryer (2005d), I provide data for both definite articles (markers) which are distinct from demonstratives and definite articles which are identical to demonstratives. The data in this paper is based solely on the former.
Fig. 14 shows the proportion of NDef order among OV languages. In this case, although the proportion for Africa (81%) is higher than the world average (69%), it is not significantly higher and there are two areas with higher proportions of 86%.

2. Distribution within Africa

The previous section shows that in most cases, postnominal position is more common in Africa than in the other five areas or is as common in Africa as one or two other areas and in only three cases does any other area exhibit a proportion that is higher than Africa by more than 1%. Furthermore, in all cases, Africa exhibits a proportion that is higher than the world average and in most cases considerably higher. We can conclude that Africa does indeed exhibit a greater preference for postnominal position of modifiers than other areas of the world.

However, the data presented so far does not tell us how broadly this preference is found within Africa. The data is consistent with its being found only with a subset of languages of Africa. In this section, we address the question of the breadth of this tendency within Africa. To do this we will examine bar graphs showing proportions similar to those in the previous section, but ones showing the proportions of genera within the traditional four language families of Greenberg (1963)^6.

2.1 Order of adjective and noun in Africa

---

^6 The use of these four sets of languages does not presuppose the validity of these four families. The status of both Khoisan and Nilo-Saharan is open to question. For the purposes of this paper, “Khoisan” can just as easily be interpreted as non-Niger Congo languages of southern Africa. Similarly, “Nilo-Saharan” can just as easily be construed as non-Niger-Congo non-Afroasiatic languages of north-central Africa, except that the language isolates Bangime and Laal are somewhat arbitrarily included here within Niger-Congo. What is important for present purposes is to have some division of the languages of Africa in order to examine the distribution and the four families of Greenberg (1963) probably serve better than most other possibilities, regardless of their genealogical validity. For convenience, I will refer to these four groups as families.
Fig. 15 shows the proportions of genera containing NAdj languages among each of the four families. It shows that the proportion is higher in all four families than the world average of 62% (shown above in Fig. 1), though it is only marginally higher in Afroasiatic.

Map 9 shows the distribution of the two orders of adjective and noun in Africa. The four families are distinguished on this and all remaining maps by the shape of the dots: squares for Afroasiatic, diamonds for Nilo-Saharan, circles for Niger-Congo, and triangles for Khoisan. These are not all distinguishable, but in at least some cases are useful. Three areas where the less common AdjN order stands out are (1) among Ubangi languages and a few Nilo-Saharan languages nearby in central Africa; (2) among some Afroasiatic languages in Ethiopia and Eritrea (including South Semitic, Cushitic and Omotic languages); and (3) in four Khoisan languages in southern Africa (all of them Khoe languages).
2.2 Order of demonstrative and noun in Africa

Fig. 16 shows the distribution of NDem order among the four families in Africa. Again, in all four families, the proportion is higher than the world average of 37%, shown above in Fig. 4. The proportion is highest in Niger-Congo and Nilo-Saharan.
Map 10 shows the distribution of the two orders of demonstrative and noun in Africa. There is little pattern to the distribution of the less common DemN type, it being fairly evenly spread throughout the continent.
2.3 Order of numeral and noun in Africa

For the order of numeral and noun, it is worth distinguishing VO and OV languages, since, as noted above, Africa is so strikingly different from the rest of the world in that most VO languages of sub-Saharan Africa are NNum, while elsewhere in the world VO languages are predominantly NumN. Fig. 17 shows the distribution of NNum order in VO languages among the four families in Africa. All four families show a proportion of NNum order that is much higher than the world average of 31% for VO languages. Furthermore, the 63% figure for Afroasiatic is lower than the other three families due only to Berber, Semitic and Egyptian-Coptic. If we exclude those three subfamilies, and include only the genera in sub-Saharan Africa, all within Chadic, the proportion for Afroasiatic goes up to 100%. In that case, all four families exhibit a proportion higher than 90%, contrasting sharply with the world average of 31% shown
above in Fig. 8 (and the world average excluding Africa is only 20%). We thus have a particularly striking difference between Africa and the rest of the world that is shared by all four families within sub-Saharan Africa.

Map 11 shows the distribution of the two orders in Africa. There are, in fact, only five VO&NumN languages in sub-Saharan Africa in my sample: Wolof, Kresh, and three Cross River languages in Nigeria (Kana, Gokana, and Ogbronuagum).
Fig. 18 shows the distribution of the two orders of numeral and noun in the four families for OV languages. Here we find that two of the families, Khoisan and Afroasiatic, exhibit proportions that are slightly lower than the world average of 53% for OV languages (shown above in Fig. 9). The proportions for Niger-Congo and Nilo-Saharan, however, are much higher than the world average.
When we examine Map 12, we see that the OV&NumN languages are found almost exclusively in two areas: (1) southern Africa (among Khoe languages) and (2) in the area around Ethiopia. The two orders of numeral and noun are about equally common among Afroasiatic languages in the area around Ethiopia and there is a rough trend for the more northern languages to be NumN and the more southern languages to be NNum: in my data, South Semitic, North Omotic, Central Cushitic and some Eastern Cushitic languages are NumN, while South Omotic, Southern Cushitic and other Eastern Cushitic languages are NNum. Apart from these Khoisan and Afroasiatic languages, there are only two other OV&NumN languages in Africa in my database. One, Shabo, is a Nilo-Saharan language spoken in Ethiopia. The other is Ijo, spoken in Nigeria. Ijo is thus the only OV&NumN language in Africa outside the two areas mentioned.
2.4 Order of genitive and noun in Africa

As discussed above, Africa contains significantly more of the exceptional OV&NGen languages than are found elsewhere in the world, where the average for Africa is 31%, much higher than the world average of 9% shown above in Fig. 10. Fig. 19 shows that this high average is found in three of the four families, the only exception being Khoisan.
Fig. 19
Proportions of genera containing NGen languages among OV languages

Map 13 shows that these OV&NGen languages are widely spread out and in this case include a number of Afroasiatic languages of the area around Ethiopia.

Map 13
Order of Genitive and Noun in OV Languages in Africa
Fig. 20 shows that the same three families, the three other than Khoisan, show a higher proportion of SVO&NGen languages than the world average of 51% shown above in Fig. 11 above.

Map 14 shows that the minority type of SVO&GenN languages are primarily found in two areas, among Gur, Kru, and Kwa languages in west Africa and among various Khoisan groups in southern Africa (Ju, ǂHoan, !Ui and Taa).
2.5 Order of relative clause and noun in Africa

Fig. 21 shows the proportion of genera containing languages with prenominal relative clauses among OV languages. Only Niger-Congo and Nilo-Saharan are significantly higher than the world average of 53% shown above in Fig. 12. Inspection of Map 7 above shows that again the OV&RelN languages, the ones with prenominal modifiers, are primarily in the area around Ethiopia and in southern Africa. The only OV&RelN language outside these two areas is Ijo, in Nigeria.
2.6 Order of definite marker and noun in Africa

Turning finally to the order of definite marker and noun, Fig. 22 shows that the three families other than Khoisan exhibit significantly higher proportions of NDef order among VO languages than the world average of 40% shown above in Fig. 13.

Fig. 23 shows that among OV languages, the three families other than Afroasiatic are exclusively NDef, contrasting with the world average of 69% shown above in Fig. 14. My data contains three OV&DefN languages and all three are Afroasiatic languages spoken in Ethiopia or Eritrea.

3. Summary

We have seen that nearly all categories of modifiers follow the noun more often languages in Africa than among languages elsewhere in the world. This trend is particularly well pronounced among Niger-Congo, Nilo-Saharan and Chadic languages. The exceptions were found most often among Khoisan languages and among the OV
Afroasiatic languages of the area around Ethiopia. However, among Khoisan languages, the exceptions are primarily among Khoe languages. Apart from consistently employing GenN order, all other Khoisan languages (Sandawe, Ju, ‡Hoan, ‡Ui and Taa languages) generally place modifiers after nouns. Among the OV Afroasiatic languages in the area around Ethiopia (South Semitic, Cushitic and Omotic), one finds a complex mixture of different word order types as far as nominal modifiers are concerned. Within Omotic, North Omotic tends to employ prenominal modifiers while South Omotic tends to employ postnominal modifiers. The pattern within Cushitic is more complex, though the most southern Cushitic languages, in Kenya and Tanzania, employ postnominal modifiers. The Afroasiatic languages north of sub-Saharan Africa (Berber, Egyptian-Coptic, and Semitic) generally conform to the trend towards postnominal modifiers, except for the most dramatic pattern, that of NNum order among VO languages.

Can we conclude from the evidence in this paper that Africa as a whole forms a linguistic area? There are two considerations that should make us hesitate from drawing such a conclusion. First, the possibility of coincidence should not be underestimated. Consider for example the fact that the VO Khoisan languages are NNum, like the VO Niger-Congo, Nilo-Saharan and Chadic languages. Is it necessarily the case that this property of the VO Khoisan languages is causally related to this property of the other VO languages in Africa? Consider the fact that, ignoring the possibility of languages allowing both orders of numeral and noun (which does happen), there are two possibilities, NumN and NNum. The possibility that the VO Khoisan languages are NNum quite independently of the other African VO languages being NNum simply by coincidence is thus quite high. The same issue of the possibility of coincidence arises with much of the overall noun-modifier order pattern I have documented in this paper. Perhaps it is a coincidence that part of Nilo-Saharan shares this property. Perhaps it is a coincidence that the Afroasiatic languages north of sub-Saharan Africa share this property.

The second consideration is that in order to demonstrate a linguistic area, one must demonstrate that a number of independent typological features are found throughout that area more often than elsewhere in the world. In effect, this is what is necessary in order to address the problem of the possibility of coincidence: the more independent typological features one can document, the less likely that any one of them is a coincidence. While Greenberg (1959), Heine and Leyew (2007), and Creissels, Dimmendaal, Frjzynogier and König (2007) have put forth claims or proposals regarding other features of Africa as a linguistic area, it is necessary to conduct detailed surveys of such features, not only within Africa, but also elsewhere in the world, in order to determine that they are more frequent in Africa, and it is necessary to show that these features are found throughout Africa or most of Africa, not just among some subset of the languages of Africa. The evidence provided by Heine and Leyew is potentially the most relevant in that it provides quantitative data from 99 African languages and 50 non-African languages, but suffers in a number of respects. First, the majority of the languages of Africa in their sample are Niger-Congo, while it contains only 15 Nilo-Saharan languages and only 6 Khoisan languages. This bias would not be so serious if they examined the distribution of the properties they discuss over the different families. But this is the second problem: all they provide is the percentage of the 99 languages that
exhibit each of the eleven properties they discuss. What this means is that unless the percentage is over 85%, we have no way of knowing that the property is widespread in Africa, as opposed to being confined to two families. What is needed for each property they discuss is to see its precise geographical and genealogical distribution. Some of the properties they discuss are clearly not widespread in Africa, such as labial-velar stops, which Clements and Rialland (2007) report are restricted to what they call the Sudanic Belt, one of only six zones they discuss (though one with a particular large number of languages).

One merit of Heine and Leyew’s study is that they include 50 languages outside Africa and compare the languages of Africa to those outside of Africa. Again, however, the only figure they provide us with is the average number of their properties that the languages in their sample in each continent have. While the average for Africa, 6.8, is twice the next highest average, 3.4, this does not tell us whether Africa is a linguistic area, since features that are widespread within Niger-Congo and some adjacent languages or within one large area could account for this. Note that the inclusion of labial-velar stops, found in 39 of the 99 languages from Africa in Heine and Leyew’s sample, pushes up the average number of properties found in languages in Africa by 0.4, but is restricted to one region within Africa. In comparing Africa to the rest of the world, we need to do this for each property, not just in terms of the average number of properties.

If some of Heine and Leyew’s properties (or some other ones) could be shown to have a widespread distribution in Africa similar to that of postnominal modifiers as demonstrated in this paper then we would have better evidence that Africa is a linguistic area and better evidence that the widespread distribution of postnominal modifiers in Africa is not simply a coincidence. However, I should note that one of Heine and Leyew’s eleven properties is the one discussed in this paper, postnominal modifiers. But it is the property found in the highest number of the languages in their sample, 89, suggesting that perhaps there is no property as widespread in Africa as postnominal modifiers. The next most common property, *surpass*-comparative constructions, was found in only 82 languages. And in the data in Stassen (2008), although the map shows *surpass*-comparative constructions to be considerably more common in Africa than elsewhere in the world, this is primarily due to Niger-Congo languages. Only three of the seven Nilo-Saharan languages in Stassen’s sample have *surpass*-comparative constructions, and the single Khoisan language does not, although both Chadic languages in his sample do.

The next most frequent property in Africa in Heine and Leyew’s data is tone, found in 80 languages. This property is in fact a better candidate for a property characterizing all of sub-Saharan Africa: in Maddieson (2008), it is found overwhelmingly in all four families in sub-Saharan Africa, while elsewhere in the world it is found in concentrated areas that are smaller than sub-Saharan Africa, such as southeast Asia and Mesoamerica. The distribution of tone in Africa is similar to that of postnominal modifiers except that it excludes the languages north of sub-Saharan Africa and includes Khoisan languages and most languages of the area around Ethiopia. We thus have at least one other property that would appear to argue for Africa, or at least sub-Saharan Africa, as a linguistic area. However, even two properties, postnominal modifiers and tone could coincidentally occur in Khoisan languages independently of
their occurrence elsewhere in Africa. Only if we find evidence of additional properties with a similar widespread distribution will we have compelling evidence for Africa as a linguistic area.

References


