The historical syntax of East Cushitic: A first sketch
Mauro Tosco, Napoli

Résumé
L'article aborde la question de l'ordre des mots original en Cushitique Orientale. Un ordre verbe-final et Modificateur-Tête du syntagme est généralement supposé pour l'ensemble du Cushitique. En réalité, les langues couchitiques orientales sont très partagées dans leur syntaxe: l'ordre Modificateur-Tête est de rigueur seulement dans les langues du Plateau éthiopique et le Saho-'Afar, tandis que toutes les autres langues (Somali et Oromo en tête), même en conservant le verbe à la fin de l'énoncé, ont en principe l'ordre Tête-Modificateur.
Dans un cas comme dans l'autre, on trouve souvent une particule ou un affixe avec accord en genre/nombre avec la Tête. En partant de l'analyse de ces éléments, l'article propose pour le Proto-Cushitique-Oriental un ordre Tête-Modificateur (sauf que le numeral était la Tête du syntagme, comme encore aujourd'hui en Somali, et que la position du démonstratif reste incertaine) et verbe-final. Dans les langues du Plateau et le Saho-'Afar cet ordre a été transformé en Modificateur-Tête, en atteignant donc une syntaxe nominale en harmonie avec la position finale du verbe. Dans les autres langues, au contraire, l'ordre Tête-Modificateur a été préservé et systématisé (en transformant, entre autres, le numeral en Modificateur).
Ce développement a été bouleversé en quelques langues par l'influence extérieure: c'est le cas notamment du Baïso, qui est en train de changer sa syntaxe originale en l'adaptant à celle des langues à ordre Modificateur-Tête par lesquelles il est complètement entouré.

0. Introduction

The status of Ethiopia as a language area has since long been well established, and Emeneau's definition of the Indian case well applies to Ethiopia: "an area which includes languages belonging to more than one family but showing traits in common which are found not to belong to the other members of (at least) one of the families" (Emeneau 1956:16, fn. 28; emphasis ours). India and Ethiopia further share the character of being linguistically born out of the encounter between one or more autochthonous language families and an originally foreign language group, the latter absorbing various traits typical of the former during the process of nativization. In

---

1 This paper has been written during my stay at the Institut für Afrikantistik of the University of Cologne as part of a project on "Grammaticalization and Reanalysis in the Ethiopian Language Area" financed by the Alexander von Humboldt Foundation.
the case of India, we have Aryan languages taking over Dravidic traits, while in Ethiopia we have the meeting of Semitic languages with, above all, Cushitic languages (both being members of the Afroasiatic phylum). This means that, while the foreign languages adapt to the autochthonous linguistic type, far less change is implied on the part of the "native" languages. One could even say that this kind of language area is defined on the basis of the characteristics of a language family as spoken outside the area in question, plus the characteristics of the autochthonous languages. In the case of Ethiopia, we have the very non-Semitic characteristics of the Ethiosemitic languages, and - knowing how a "good" Semitic language must look like - we can explain the characteristics in terms of those of the autochthonous language family. In a classical interpretative scheme, the linguistic history of Ethiopia was thus seen as the result of an imperfect shift of Cushitic-speaking populations to a superior Semitic language and culture. Although this position is untenable today, there is no doubt that the main traits of the modern Ethiosemitic languages are to be ascribed to a Cushitic substratum, to be identified more specifically with languages of the Agaw (Central Cushitic) group together with, for the Southern Ethiosemitic languages, "Sidamo" - i.e., in modern terminology, the Highland East Cushitic group of languages. The central role - in both geographical and historical terms - of these languages has certainly contributed to ascribing to Cushitic languages as a whole certain syntactic traits which are actually shared only by a minority of the Cushitic languages.

Word order plays a decisive role in the identification of the Ethiopian language area: the SOV basic word order of most languages of this area is the very first grammatical feature listed by Ferguson (1976).

Actually, "consistent" V-final languages are but a minority: in a large part of the Cushitic-speaking area the normal position of nominal modifiers is after their head, even though the verb retains the last position in the sentence and postpositions rather than prepositions are found. Languages of this type are the Omo-Tana languages (Arbore, Elmolo, Dasenech, Bayo, Somali - including Boni - and Rendille), the Omotic languages (Konso, Dirayta or Gidoles) and the Dullay cluster.

In Black's (1974) classification, all these languages except Dullay (Black's "Werizoid") are grouped together into the South Lowland East Cushitic group, which, together with Saho-Afar, is one of the two branches into which East Cushitic is divided, the other one being Highland East Cushitic (HEC, made up, from North to South, of Hadiya, Kambata, Sidamo, Gedeo - or Darasa - and Burgi).²

² The following abbreviations will be used:

| ABS | Absolute   | Mod | Modifier   |
| Adj | Adjective  | N   | Noun       |
| Art | Article    | NOM | Subject form of nominals |
| Cn  | Connector  | Num | Numerical  |
| Dem | Demonstrative | P-EC | Proto-East Cushitic |
| Det | Determinate | P-HEC | Proto-Highland East Cushitic |
| EC  | East Cushitic | Pl | Plural |
| F   | Feminine   | P-LEC | Proto-Lowland East |
| Gen | Nominal Genitive | | Cushitic (minus Saho-Afar) |
| GM  | Gender marker | Poss | Pronominal Genitive |
| HEC | Highland East Cushitic | Pro | Pronoun |
| LEC | Lowland East Cushitic | Rel | Relative clause |
| M   | Masculine | S | Singular |

The original transcription is followed throughout, for Somali the national orthography is used.
odds with a "pure" V-final syntax. Although the authors consider the EC languages to have been more or less strictly V-final historically, they point out that, conversely, the V-final HEC languages have deeply borrowed from the consistent SOV Omeno languages (North Omotic).

In this paper I shall argue that the strict V-final type normally associated with Cushitic needs reexamination. Proto-EC (P-EC) will no longer represent a typologically consistent stage from which the modern languages have departed, generally under external influence. Rather, in line with the uniformitarian principle of diachronic typology, proto-EC can be ascribed the same amount of typological inconsistency which is found in the modern languages (and in most of the world's languages, for that matter). The following table (Table 1) shows the basic word order patterns of representative East Cushitic languages. All languages having SOV as basic word order and postpositions, these traits are not included. With "Gen" nominal genitive is meant, pronominal genitive is labelled "Poss". The two rightmost columns give the number of pre- and postmodifying constructions; bracketed numbers indicate that a construction alternates with, or is not basic compared to, another one; "Mod" (Modifier) is used as a cover term for all dependent categories. When the Mod is affixed to its Head, "+" is written between the two.

<table>
<thead>
<tr>
<th>Language</th>
<th>Adj, N</th>
<th>Gen, N</th>
<th>N, Poss</th>
<th>Dem, N</th>
<th>Num, N</th>
<th>Rel, N</th>
<th>No. of Mod N NP s</th>
<th>No. of N Mod NP s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HEC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hadiya</td>
<td>Adj N</td>
<td>Gen N</td>
<td>Poss+N</td>
<td>Dem N</td>
<td>Num N</td>
<td>Rel N</td>
<td>6</td>
<td>Ø</td>
</tr>
<tr>
<td>Kambata</td>
<td>Adj N</td>
<td>Gen N</td>
<td>N+Poss</td>
<td>Dem N</td>
<td>Num N</td>
<td>Rel N</td>
<td>6 (1)</td>
<td></td>
</tr>
<tr>
<td>Sidamo</td>
<td>Adj N</td>
<td>Gen N</td>
<td>N+Poss</td>
<td>Dem N</td>
<td>Num N</td>
<td>Rel N</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Gedeo</td>
<td>Adj N</td>
<td>Gen N</td>
<td>Poss N</td>
<td>Dem N</td>
<td>Num N</td>
<td>Rel N</td>
<td>6</td>
<td>Ø</td>
</tr>
<tr>
<td>Burji</td>
<td>Adj N</td>
<td>Gen N</td>
<td>Poss N</td>
<td>Dem N</td>
<td>Num N</td>
<td>Rel N</td>
<td>6</td>
<td>Ø</td>
</tr>
<tr>
<td>&quot;LEC&quot;</td>
<td>Adj N</td>
<td>Gen N</td>
<td>Poss N</td>
<td>Dem N</td>
<td>Num N</td>
<td>Rel N</td>
<td>6</td>
<td>Ø</td>
</tr>
<tr>
<td>Dallay</td>
<td>N Adj</td>
<td>N Gen</td>
<td>N+Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>Ø</td>
<td>6</td>
</tr>
<tr>
<td>Gomoid</td>
<td>N Adj</td>
<td>N Gen</td>
<td>N Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>Ø</td>
<td>6</td>
</tr>
<tr>
<td>Omo/Tana</td>
<td>N Adj</td>
<td>(N Gen, (Gen N+Poss))</td>
<td>N+Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>1 (2)</td>
<td>5</td>
</tr>
<tr>
<td>Somali</td>
<td>N Adj</td>
<td>N Gen</td>
<td>N Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>Ø</td>
<td>6</td>
</tr>
<tr>
<td>Arbores</td>
<td>N Adj</td>
<td>N Gen</td>
<td>N Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Bayso</td>
<td>N Adj</td>
<td>N Gen</td>
<td>N Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>Ø</td>
<td>6</td>
</tr>
<tr>
<td>Dasenese</td>
<td>N Adj</td>
<td>N Gen</td>
<td>N Poss</td>
<td>N Dem</td>
<td>N Num</td>
<td>N Rel</td>
<td>Ø</td>
<td>6</td>
</tr>
</tbody>
</table>

When we look at the internal structure of the NP.s, we find various elements between the Head and (at least some of) the Mod.s in both postmodifying and premodifying languages. My hypothesis will be that the presence of these elements is a clue to a historical change of the order of the elements within the NP.

1. Gender-sensitive suffixes in HEC

In all HEC languages a few Mod.s show various suffixes, often agreeing with the Head in gender/number, and glossed in the following "GM" (gender marker). Their historical origin and synchronic function are, nevertheless, more complex, as will become clear from the following discussion.

Table 2 summarizes the GM.s on Mod.s in HEC languages.

---

Table 2
Gender-sensitive suffixes in HEC

<table>
<thead>
<tr>
<th>Language</th>
<th>V/Rel</th>
<th>+</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadiya</td>
<td></td>
<td>-ha</td>
<td>N</td>
</tr>
<tr>
<td>Kambata</td>
<td></td>
<td>-ta</td>
<td>N</td>
</tr>
<tr>
<td>Sidamo</td>
<td></td>
<td>-u</td>
<td>N</td>
</tr>
<tr>
<td>Dallay</td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Gomoid</td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Omo/Tana</td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Somali</td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Arbores</td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Bayso</td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Dasenese</td>
<td></td>
<td></td>
<td>N</td>
</tr>
</tbody>
</table>
The common denominator seems to be that whenever only one Mod is gender-marked, this is the Rel; when two Mods are marked, they are the Rel and the Gen. This is, of course, consonant with the marked character of proposed Rel.s which, as is well known, often follow their head N even in strictly V-final languages. The Poss Mod stands apart and has different constructions.

1.1. Gender-marked genitive, possessive and relative phrases

Hudson (1976:260), who calls these suffixes "Articles", reports for Hadiya, the northernmost HEC language, an optional and invariable -k suffix on the verb of the Rel clause, especially with "noun clauses based on the relative" (i.e., nominalized Rel.s). For Kambata, the same author notes that "the endings -h/-t are common with the nominalized relative clause, but optional otherwise". In Sidamo there are "three relative pronouns: masculine -h, feminine -t, but also -r, for the plural common". More widespread is the use of these forms in the southern HEC languages Gedeo and Burji where, especially in the latter, they are also used as "true articles". In Gedeo, -k/-t (followed by -i in the examples) are "rarely used with nouns in undefined oblique case environments [...] there is with the relative clause [...] a pair -k/-tt". In Burji, -k (Hudson 1976 has -g and -h) and -i are used with definite nouns as well as Rel clauses.

When we turn to the expression of the Gen phrase, again we find no suffix in Hadiya, except for the glottalization of a word-final sonorant in the genitive (Hudson 1976:254). Plazikowsky-Brauner (1960:44f.) noted that the Gen is marked by the dropping of the final vowel of the N, coupled, for sonorant-final N.s, with the insertion of /r/ before the sonorant, together with the changes /lr/ > /pr/ and /fr/ > /fr/. N.s with the -ro formative and /r/-final N.s add -r "das aber labil und abfallen kann".

For Kambata, Cerulli (1925:648) says that "il nome, reggente o retto, su cui si vuol richiamare l'attenzione ha il suffisso determinativo -u, -o, o, più raramente, -i; the same suffixes are used "quasi sempre" with the Gen, as in danni bieto 'la casa del giudice', and mini manada 'la gente della casa'.

This does not appear in Hudson's (1976:254) examples of Gen constructions in Kambata, but one must note that all his examples involve proper names in Gen position. Leslau (1952:352) notes that vowel-final N.s change their final vowel to /u/ or /o/, while F N.s in -ta drop this element.

Sidamo, Gedeo, and Burji have special suffixes for the Gen, different from the ones found with the Rel. In Sidamo, -u (M) and -te (F) are reported by Hudson. According to Cerulli (1938:123), -u and -te are found not only on the Gen but also on the Adj, Num, and in PP.s, provided the possessor is not a proper N:

bu'u mana 'uomini cattivi' (bu'u 'cattivo')
susu bieto 'tre razzisti' (sasu 'tre')
dubbā giddu 'dentro il bosco' (dubbū 'bosco')

In both Gedeo and Burji, the suffixes found with the Rel are preceded in the Gen by an infix -n- which is absent from proper N.s:

Gedeo: -nka (M)/ -nt'a (F) (Abs.) and -nki (M)/ -nt'i (F) (Nom);
Burji: -nka (M)/ -nts (F) (Abs.) and -nku (M)/ -nts (F) (Nom).

When we compare the realization of Rel and Gen NP.s (for non-proper N.s) in HEC languages we get the following picture (Table 3):
Table 3  
N, Gen and N, Rel in HEC

<table>
<thead>
<tr>
<th>Gen, N</th>
<th>Rel, N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadiya</td>
<td>Gen (+ Det ?) N</td>
</tr>
<tr>
<td>Kambata</td>
<td>Gen + Det (M/F ?) N</td>
</tr>
<tr>
<td>Sidamo</td>
<td>Gen + Det, M/F N</td>
</tr>
<tr>
<td>Dirasa</td>
<td>Gen + -n + Det, M/F N</td>
</tr>
<tr>
<td>Burji</td>
<td>Gen + -n + Det, M/F N</td>
</tr>
</tbody>
</table>

(in Sidamo, Det, stands for the -u (M) and -te (F) set)

Consider now the expression of a pronominal genitive phrase in HEC (Table 4):

Table 4  
Possessive phrases in HEC

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadiya</td>
<td>Poss + N</td>
</tr>
<tr>
<td>Kambata</td>
<td>N + Poss ~ Pro + Det, M/F N</td>
</tr>
<tr>
<td>Sidamo</td>
<td>N + Poss</td>
</tr>
<tr>
<td>Gedeo</td>
<td>Pro + Det, M/F N</td>
</tr>
<tr>
<td>Burji</td>
<td>Pro + Det, M/F N</td>
</tr>
</tbody>
</table>

In the southern languages, suffixed Poss.s are not found and the Poss.s are overtly (in Gedeo, less so in Burji) made up of independent Pro.s (the NOM forms in Gedeo, the ABS forms in Burji) plus the gender markers. Kambata too has prefixed Poss.s, again made up of ABS Pro.s plus the determiners. This construction alternates with suffixed Poss.s, which is the only possibility in Sidamo.

Typologically, Poss. N is the expected order in these languages, but anywhere except in Hadiya this order is morphologically more complex than the alternative N + Poss construction of Sidamo and Kambata; the latter, moreover, has bound Poss rather than free forms as in the premodifying order. The premodifying forms of Gedeo, Burji and, partially, of Kambata, can be considered innovations as they are more elaborate morphologically and follow the same pattern as the Gen and Rel constructions.

Only Hadiya is consistently V-final, having prefixed Poss; the other languages have either suffixes (Sidamo, optionally also Kambata), or their prefixed Poss.s are clearly derived from the independent Pro.s plus the gender-sensitive markers found with Gen (Gedeo, Burji, and, as an alternative to N + Poss, Kambata). Even Hadiya, which has eliminated the suffixes on Gen (except for the glottalization, see above) and Poss. (if it ever had them), has Poss.s clearly derived from the independent Pro.s. More irregular are the Poss suffixes of Sidamo in comparison with their ABS Pro.s counterparts, especially in 1S and 2S:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadiya</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sidamo</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Moreover, the suffixed Poss. of neither Sidamo nor Kambata show any trace of determiners before the Poss (Kambata has gender-sensitive -h-/t-after the Poss). The presence of these suffixed Poss. is coupled with another very irregular trait – in terms of a consistent V-final language – i.e., the presence of suffixed object Pro.s. One can speculate that the latter have helped the Poss affixes to keep their original postmodifying position and that N + Poss is the original order, while Poss + Det N is an innovation linked with the presence of premodifying Gen and Rel.

1.2. Copulas and the historical development of gender markers in HEC

We propose to ascribe the presence of GM.s on the Mod.s in HEC to the left-dislocation of a topical Mod: In this construction, the original value of the GM itself was probably that of a copula element which followed the left-dislocated Mod – the whole construction being therefore a cleft. When we look at the realization of the copula in HEC languages, we find at least partial support for this reconstruction:

Table 5  
Copulas in HEC

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadiya</td>
<td>-tte (opt. for M; otherwise Ø)</td>
<td></td>
</tr>
<tr>
<td>Kambata</td>
<td>-[h]a (M), -ta (F); Ø if GM already present; -ti for Pro.s and human N.s</td>
<td></td>
</tr>
<tr>
<td>Sidamo</td>
<td>-u (M), -te (F); -ti for Pro.s and human N.s</td>
<td></td>
</tr>
<tr>
<td>Gedeo</td>
<td>-ke-n (M), -te-n (F, Pro.s and human N.s); -n if GM already present</td>
<td></td>
</tr>
<tr>
<td>Burji</td>
<td>-na- -k'</td>
<td></td>
</tr>
</tbody>
</table>

Copulas derived from the gender markers k and t (*k > h in Kambata and *K > Ø in Sidamo; *k > k’ in Burji is unexplained) whose presence is
linked to the determinate status of the subject of the clause are found in three out of the five HEC languages, being absent only in the geographically extreme Hadiya and Burji.

We find in HEC languages two other copular elements: a second element based on t and, in the southern HEC languages Gedeo and Burji, a copula based on n, coupled with the GMs in Gedeo (if the N is [+Det]) and in (free?) alternation with k’a in Burji.

The proviso that the N must be [+Det] for the gender markers to appear as copulas is accounted for if we bear in mind that the GM in these languages is only found on the Gen and the verb of (nominalized) Rel.s; therefore, the absence of the copula when the GM is found on the Gen is simply due to the fact that the copula is already there in the Gen+GM construction (which is, at least diachronically, a cleft construction), and a further copula is therefore excluded.

No such restriction applies to the copula based on *t of Kambata and Sidamo; for the latter language, a te (for F [+Det]-N.s) ~ ti alternation is mentioned by Cerulli (1938:97), who reports the use of te as a copula after a Rel clause, translating the whole with a cleft construction, as in:

\[ attì mančó ga’á qaqqó itera te \]
\[(lit.: “this is the woman that will tomorrow ...”)\]

"questa donna domani partorirà un figlio"

and

\[ ninke losasinare ti \]
\[(lit.: “this is us who are working”)\]

"noi lavoriamo"

While for Cerulli te and ti are in free alternation, at least in these two sentences the pattern seems to be: te with a F N (mančó), ti with a Pro. Thus, Sidamo te goes with Kambata -ta and Gedeo -te-a, while ti goes with its homophonous counterpart in Kambata about which Leslau (1952:353) observes that it is used in cleft constructions, though without providing examples.

As for the ta/te ~ ti alternation in Kambata and Sidamo, the first is used only with F N.s, while the latter is used with Pro.s and human N.s. While only Kambata and Sidamo have a special ti copula for Pro.s, in all HEC languages the personal Pro.s (except of course 3MS) are treated as F: in Burji (Hayward 1988:687), Gedeo (which has -te-a with Pro.s), and Hadiya (for which Hudson [1976:274] reports anette ‘it’s me’). This is therefore a pan-HEC trait.

It seems possible to reconstruct *te as an invariant copula, originally used with all referentially determinate N.s and with Pro.s, and not restricted to F.

The origin of this copula can be sought in the particularizing suffix -tV widely found in Cushitic. As to the copula used with indeterminate N.s, it was probably just Ø.

Not only was *-tV not restricted to F N.s, but we also find evidence that its counterpart *k was similarly not restricted to M N.s — witness the optional use of -k after Rel.s in Hadiya, as in letháʔo-t-k ‘she who played’ (Hudson 1976:260). Note further that Hadiya has no specific copula for M N.s, -te being, at least optionally, used with all N.s; therefore, -k and -te are in complementary distribution. On the other hand, *k as an unrestricted copula is found in Burji as k’a.

In the central HEC languages Kambata and Sidamo *kV and *tV were associated with M and F gender, respectively, and in their copular function they followed the Gen in its left-dislocation position before the head. Although clearly gender markers, Sidamo and Kambata *kV and *tV derivatives still keep their original determinative value.

The original *tV copula was preserved as ti used with Pro.s and in cases in which a human N is not yet determinate (i.e., whenever a gender-sensitive copula is not yet found on the Mod): this is the situation represented by Sidamo -u, -te vs. -ti and Kambata -[h]a, -ta vs. -ti. In the southern languages Gedeo and Burji the determinative value of *kV and *tV was lost, and these languages had recourse to a "new" determinative element *-n(V). In left-dislocation of the Gen, this element preceded the gender markers, yielding Burji -nka, -nta, Gedeo -nka, -nta, etc. The determinative value of n- is evident from its absence from inherently determinate N.s, such as proper N.s. That *kV and *tV derivatives no longer have determinative value is apparent from their use as indeterminate Nominative markers in Burji (cf. Hayward [1988]); it is this use that led Hudson [1976] to interpret these elements as "articles".

No -n element is found with Rel.s, either, possibly because restrictive Rel.s are inherently determinate, and the use of the simple gender-sensitive copula marked left-dislocation; we thus have Burji -gu (< *-ku) and -[i] (< *-fi), and Gedeo -ke, -te.

In its turn, -n too became a copula, following the gender-markers in Gedeo and coupling with -k’a in Burji. The use of -k-t as "connectors" linking the Head and the Mod.s is of course widespread outside HEC. Apart from the connectors of LEC languages (dealt with in 2.), we can mention the focus markers of Dallay of which -kka (or k-prefixed to the preverbal selector) yields a contrastive focus, while the element ta after the N (or t- prefixed to the verbal selector) is a more rarely used "adjective" focus marker (Amborn, Minker, Sasse 1980:82-83). The origin of the use of these elements as focus markers seems to be found most naturally in...
1.3. A note on 'Afar

Within EC, Saho-'Afar is the only strictly premodifying language outside the HEC group. The normal order for both Gen and Poss is Gen N, Poss N (with Gen marked by case endings and the Poss series being derived from the ABS series). Alternatively, it is possible to use the postposition -k. In this case, Gen+k N is found with nominal Gen.s, whereas with Poss the normal order is N Poss+k: amó yok (head me+k) "my head" (Bliese 1981:171). When -k is used, Gen.s do not have the "genitive" case; likewise, with Poss the ABS forms of the Pro.s are used rather than the GEN.

The only possible explanation for the N Poss order is historical, as a relic of a once more widespread postmodifying syntax. Black (1974:239), too, assumes possessive suffixes for his PO-LEC and considers the preposed affixes of Saho-'Afar an innovation.

1.4. Preliminary conclusions on word order change in HEC

The proposed reconstruction of the development of the copula in HEC allows us to analyze the Mod N order, as far as the Gen, Rel and Poss Mod.s are concerned as the result of the grammaticalization of a cleft construction in which the copula is suffixed to the cleft in conformity with the final position of the verb, and being moreover formally similar to a gender marker - was reinterpreted as an agreement marker.

Although the N Mod order is found today only in a part of HEC in the suffixed Poss of Sidamo and Kambata, and nowhere in the Gen and Rel, it is possible to reconstruct for Proto-HEC (P-HEC) a N Mod word order for all three Mod.s.

When we turn to LEC languages, we not only find stronger evidence for this reconstruction, but also clues as to the original order of the N and the "lighter" modifiers Adj, Num, and Dem.

2. Connectors in LEC

In many LEC languages most Mod.s follow their Head, with an interposed element whose function is to "link" the preceding Head with its Mod(s). This morpheme, usually called "Connector", is often gender-sensitive. Although it has no determinative value per se, its nearest cognates are the determiners of the Sam languages (Somali dialects — including Boni — and Rendille) which are made up of two elements: the gender markers k and t (or one of their allomorphs) and various, mainly vocalic, elements with a wide range of meanings ("true" article, anaphoric, possessive, etc.; this is their use as copulas, thus paralleling the HEC situation. 

*Among the Oromoid languages, only Konso is here discussed; Somali and Rendille are excluded (given the secondary character of the determiners system), likewise Elmo. As pointed out by Sasse (1991:76), what we know of Elmo — today extinct — from Heine (1975/6) was probably obtained from a "terminal speaker" and numerous morphological and syntactic inconsistencies make the evaluation of Elmo data particularly uncertain for comparative purposes.
Dullay (Gollango dialect)

Gen, N: N \ Gen + \{ -lo (M) \} \ Gen + \{ -tal (F) \}

Adj, N: N \{ hu ~ hi (M) \} \ Adj + \{ -a (M) \}

Dem, N: N \{ l (F) \} + Dem \{ h (Pl) \}

Poss, N: N \{ t i ~ e (F) \} + Poss \ N + Poss/Kinship N.s \ h (Pl)

Rel, N: N \{ hu ~ hi (M) \} \ Rel \{ hu (Pl) \}

Arbore

Gen, N: N + \{ -l (L) \} \ a (M/ _C#) \ Gen + \{ (e)L \}

Adj, N: N + \{ -ha (M) \} \ Adj + \{ -a (M,F) \}

Dem, N: N + \{ -ha (F) \} \ Dem \{ -a (Pl) \}

Poss, N: N + \{ -ha (M) \} + Poss \{ -ha (Pl) \}

Rel, N: N + \{ -ha (M) \} \ Rel \{ -ha (Pl) \}

Dasenech

Gen, N: N \ Gen + \{ -lel (F) \} + -ka

Rel, N: N \ Rel + -ka

Adj, N: N \ Adj + -ka

Konso

Gen, N: N \ (2a) \ Gen

Poss, N: N \ Poss

Adj, N: N \ Adj \ X \ 2a \ N

Dem, N: N \ Dem

Num, N: N \ Num

Rel, N: N \ 2a \ Rel

2.1. Relative, Genitive, and Possessive

In every language except Bayso, the Rel follows the N (as it is indeed expected to do in postmodifying languages). The Cn is found in Arbore, and the article in Somali and Rendille (restrictive Rel.s always being determinate); it is also normally found in Dullay, except for "short" Rel.s which are then marked by the verb in the Subjunctive Mood (Amborn, Minker, Sasse 1980:93). The Cn is also found in Konso and Arbore, and postposed -ka in Dasenech.

The Gen of many LEC languages shares with HEC the property of being marked differently from other Mod.s. Just as in HEC the Gen N construction is morphologically more complex than, e.g., Adj N, in LEC the N Gen complex is distinguished from other NP.s by the fact that often no Connector is found between the head N and the Gen, or, although a Cn is found, the Gen has other suffixed morphological material not found on other
Mod.s. The first case (N Gen, without Cn) is found in Arbore and Dullay. The absence of the Cn has nothing to do with the "lightness" of the NP, as in the same languages the Cn is found, e.g., on Dem.s. The Cn is, on the other hand, found in Bayso, Dasenech, and in Konso (but see below).

Elmolo, too, had an element a between a non-vowel-final N and the Gen (Heine 1975/6:289) – probably not a Cn, but the same a as is found (among other allomorphs) on C-final M nouns in the closely related Arbore. In Somali, the presence of the Art as a suffix on either the N or the Gen is simply determined by the [+Det] status of the N and/or the Gen. The "genitive ending" -eed cannot be followed by the Art, with the exception of compound nouns in which the Art agrees with the head rather than the noun. The Poss.s require the presence of a Cn even in those languages where none is found with Gen.s, such as Arbore and Dullay. In many languages the Cn can be prefixed to the Poss.s which thus become phonologically autonomous pronouns; the same happens in Somali (where, of course, the Article takes the place of the Cn). In Dullay, the Cn is shunned by a few kinship terms.

A second series of GM+determinatives is found in Somali, absent, again, from a few kinship terms:

- abbahay 'my father'
- casharkayga 'my lesson'
- father+M+1S
- lesson+M+1S+1M+1Art

2.2. Adjective, Demonstrative, and Numerals

In most LEC languages, Adj.s are stative verbs. In attributive position they are formally Rel.s and pattern as such (cf. Banti 1988), being therefore postmodifying (at least in the basic word order) in all LEC languages except Bayso. The presence of the Cn in N Adj is obligatory in Arbore, normal in Dullay.

The head in Arbore N Gen constructions shows specific endings (-[Hi] for M, -e for F) which are possibly relics of case markers or, more probably, of determiners.

Whether a Cn is present or not, the Gen is marked in many LEC languages in at least a subset of Ns by an ending whose general form is *-V(G). This marker seems absent in premodifying languages. If we hypothesize that N Gen is the original order of P-LEC, just as of P-HEC, this morpheme can be interpreted as an original determiner, rather than, as is usually done, a genitive case ending. In a N Gen construction, we would indeed expect the determinate marker to appear at the end of the phrase. A clue to this original definitizing value of -V(G) is still found in Dullay where every Gen is followed by gender-sensitive vowels. The latter are absent when a further Mod follows; interestingly, a further Mod after the Gen requires the presence of the Cn.

The Historical Syntax of East Cushitic

When the verbal status of the Adj is less evident, we often find gender agreement between the Adj and its Head, as in Arbore, and also a Cn between the two. In languages with verbal Adj.s, conversely, no agreement is involved (except for the possibility of reduplication as a sign of plurality). Thus, in Somali the presence of the Art is simply linked to the [+Det] status of the head:

niinka weyn 'the big man' vs. nin weyn 'a big man'

Dem.s, too, are always postmodifying and often Cn-introduced. In Omo-Tana languages Dem.s are suffixes rather than phonologically autonomous words, and in Somali and Rendille they have entered the class of "determiners", alongside the Art, the Poss, etc.

We have already noted that Num is the only Mod that is never Cn-introduced and is also always postnominal, except in Somali. While the other languages have the harmonic order N Adj - N Num (and HEC languages the equally harmonic Adj N, Num N), in Somali the dominant order of the two Mod.s is: N Adj - Num N (see below, 3.). This idiosyncrasy is coupled with a syntactic reanalysis of the Num: in Somali, Num.s are nouns, and the N Num construction is syntactically a genitival phrase. The Num is the head of the NP, the counted item follows as a Gen, and only the Num has the Art if the whole NP is determined:

N/Num +Art / Gen/countered item

2.3. Bayso

Among LEC languages, Bayso is in many ways exceptional: it has Adj N and N Gen word order, and Rel prepoused to the N – though the tendency of Rel to be postposed even in "strict" V-final languages is well-known. Both Adj N - N Gen and Rel N - N Gen orders should be "prohibited" according to Hawkins (1983).

For both Adj and Rel alternative postmodifying orders are possible: as for Adj.s, the N Cn Adj order is used when other Mod.s are present after the N:

- dabaalja o til heelani goeran
- heifers Cn my red have-died
- "my red heifers have died"

(Hayward 1979:113)

The N Rel order is used for postpositional and temporal – i.e., heaviest – Rel.s, which require the presence of the Cn, followed in its turn by invariable ko:
2.4. Konso

Konso shows numerous word order inconsistencies, dealt with by Bliese and Gignarta (1986). Konso has a ubiquitous morpheme qa-; in Gen constructions, this marker is normally assimilated to a final /a/ of the head N but is present when a Dem follows the head.

While Konso has basic N Adj order, for most Adj.s also Adj N is possible. Whereas gender marking on the Adj is optional for N Adj word order, it is obligatory for Adj N: pórá killáá (river narrow) or killá-aáya pórá (narrow-F river) "narrow river" (Bliese and Gignarta 1986:9).

Adj N is not only morpho-syntactically marked (as it involves the obligatory gender marking of the Adj) but it is also lexically restricted: certain basic Adj.s, like most color terms, only occur in N Adj order. Thus, only kanáa pórá (boy black) "black boy" (Bliese and Gignarta 1986:10) is possible, not the reverse. Moreover, in Adj N the N is preceded by the particle qa and is tonally treated like a Gen; qa- is elided after /a/ but is found when a Dem follows the Adj and precedes the N:

kókkoon-ayoo-siné qa-ánná "those strong boys"

Just like Bayso ka/t/a/o, Konso qa- has retained pronominal value in its use in headless Rel.s:

qa-Dé "he who came ate"

who-came

Bliese and Gignarta (1986:17)

Even Adj.s can be preceded by qa- when following a Gen that is not itself preceded by qa-. While Bliese and Gignarta analyze qa- as basically a Gen marker, undergoing assimilation in most contexts, it seems possible to consider the Gen construction of Konso as simply N Gen (as in Dullay and Arbore) plus tonal modifications. The occurrence of qa- would then be required, again as in Dullay, by the presence of a further Mod (a Dem suffixed onto the N or a Gen preceding an Adj).

External pressure for Mod N order in Konso is not so easily found as for Bayso (Konso’s nearest neighbors are Oromo and Dullay, both postmodifying languages). It is interesting that the Adj is not simply preposed to the N, as in Bayso: it becomes the head of the Adj NP, with the following N preceded by qa- and having Gen features. This is reminiscent of the Num N construction of Somali in which an older word order was maintained through syntactic reanalysis. In the case of Konso, however, we have a syntactic reanalysis that makes a word order change possible without contravening the postmodifying syntactic syntax of the language.

---

7 The same rule is involved in the Gen N+Poss construction of Somali (alternative to N Gen, and particularly common in the southern dialects).
2.5. Dasenech

In Dasenech, typological irregularity is not given by the position of the Mod.s (all postmodifying), but by its invariable Subordinator -ka which is found after the Mod, instead of before it. -ka marks the end of the NP (except with Dem, Poss, and Num) in a way very natural for postmodifying languages but completely isolated within LEC. There are indications that Dasenech, too, went through a stage of Mod N word order: "-ka is dropped in some commonly used expressions which constitute a whole [...] Moreover, there are some compound nouns with the reverse order and without -ka" (Sasse 1976:205). This element can be interpreted as an old determiner, frozen and extended by analogy to N Gen (influence from a Nilo-Saharan language is suggested instead by Bliese and Gignarta 1986:32).

3. Bringing HEC and LEC together: typological and historical implications

The following table (Table 7) brings together the findings of Tables 2 and 6 concerning the presence of either the Connectors or the Gender-markers in EC:

<p>| Table 7 |
| Connectors and Gender markers in EC |
|----------|---------|----------|---------|---------|---------|</p>
<table>
<thead>
<tr>
<th>Adj, N</th>
<th>Gen, N</th>
<th>N, Poss</th>
<th>Dem, N</th>
<th>Num,N</th>
<th>Rel, N</th>
<th>no. of Cn/ GM-marked constr.s</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hadiya</td>
<td>(+)</td>
<td></td>
<td>(+)</td>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kambata</td>
<td>(+)</td>
<td></td>
<td>(+)</td>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sidamo</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Gedeo</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burji</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OROMOID</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>2 (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Konso</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DULLAY</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The absence of the Cn in Arbore and Dullay, the non-obligatoriness of the article on the head of the N Gen phrase in Somali, and the very presence of the genitive case ending -Vt in most languages point to the originality of the N Gen order. This is also the conclusion we arrived at for HEC through the analysis of the GM-suffixes. It seems therefore possible to reconstruct the order N Gen for the whole of P-EC.

The originality of N Rel is likewise beyond doubt, being the order found in every language except Bayso (which has the marked Rel N order). As to Poss, it seems possible to reconstruct the following development:
1. P-LEC had postmodifying Poss.s but, unlike their P-HEC correspondents, they were (or could also be) phonologically autonomous words when preceded by a determiner;
2. in a few languages, this possibility was lost and the Cn.s were criticizied onto the Poss, thereby losing their determinative value as in Dasenech invariable Poss.s (e.g., 1S -cu, 2S -ku);
3. in still other languages (such as Somali and Rendille), a new determinative series was suffixed to the Poss, creating the morphological sequence: N-Ad+Poss+Art.

As is well known, "lighter" Mod.s (Adj, Dem, and Num) can far more easily be moved from one side of their head to the other than Gen and Rel. This has been accounted for by Hawkins (1983:92ff.) with his "Mobility Principle", and more recently Dryer (1988) has even suggested that no typologically significant generalization about the order of Adj and N can be made.

According to Hawkins' (1983:82) Implicational Universal # VI', Dem and Num are in their turn more movable than the Adj. In principle, therefore, the absence of a Cn in N Adj constructions in LEC does not contradict the possibility that a word order change has occurred: the Adj could simply have shifted from pre- to postmodifying position without the morphological
The machinery that was required by the far heavier Gen and Rel for their opposite shift in HEC.

The originality of N Adj in EC has also been proposed by Banti (1988), who points out that this is also found among Agaw languages in Bilin. Within Oromo, the Western dialects have Adj N, but in this case the originality of N Adj order is beyond doubt. Banti considers N Adj the more conservative order within EC and thinks that the HEC languages, Agaw, etc. have adopted Adj N in accordance with their general V-final syntax. The verbal character of Adj s in LEC languages makes it reasonable to suggest that the postmodifying order assumed for Rel s was the rule for Adj s, too. For Num, it seems possible to reconstruct for PEC the Somali structure N Num Gen headed item. While this construction is far from rare among the world’s languages, the isolated position of Somali within LEC languages in this respect (even its closest relative, Rendille, has simply N Num) makes one suspect that Somali preserves here an older word order. Finally, as we noted above (1.1), the same construction is given for Sidamo by Cerulli (1938).

Note that, being in reality a N Gen construction, the position of the Num was consonant with the general postmodifying order of P-EC. As to its further demise in every language except Somali and Sidamo, one must remember that the general typological preference for harmonic patterns (like Adj N - Num N, or N Adj N Num) against disharmonic, even though dominant (like N Adj - Num N) ones (cf. Croft 1990:209). Therefore, though consonant with the syntax of the language, the position of N Num before the N was per se anomalous.

Finally, it is difficult to infer anything about the original position of Dem: it is often Cn-introduced in LEC languages, but has no GM-suffix in HEC thus pointing to an original premodifying position, which seems also preferred crosslinguistically (Dem N being, at least according to Greenberg and Hawkins, the dominant pattern).

Summing up, our reconstructed P-EC had consistent N Mod word order, except perhaps for the Dem; the Num preceded the counted item which followed as a Gen. Except for the position of Dem, this is identical with the pattern found today in Somali, and — except for the position of Num — in most modern LEC languages.

Under this assumption, the main burden of the shift fell upon HEC and Saho-Afar, while LEC languages mostly kept their postmodifying syntax, except that Num was reanalyzed in all languages except Somali as a typical Mod – and was therefore shifted after its head —, and that Dem was likewise shifted after the head, in many languages thereby losing phonological autonomy.

The same development as occurred in prehistoric times in HEC is presently witnessed in Bayso which has basic Adj N order with strong morphological traces of N Adj ( retained in more complex constructions). The shift to premodifying Adj can also be found in Konso where it is still linked to the reanalysis of Adj N as a N Gen construction (with the semantic head N taking tonal genitival features and prefixed /p/). The evidence from these typologically peripheral LEC languages is consonant with the well-known mobility of Adj s relative to their head and brings further support to our historical reconstruction of the development which led to the present consistent word order of HEC and Saho-Afar. The evidence from Bayso and Konso furthermore indicates that the word order shift was initiated and probably triggered by the Adj.

It is tempting to speculate about the cause of the HEC word order shift, eventually ascribing it to contact from the neighboring rigid SOV Omotic languages. At the present state of our knowledge, however, this seems as much of an escape-hatch as any explanation in terms of an internally-driven drift.

Certainly, the very hypothetical pressure of premodifying Dem, even coupled with Num before the counted item, was too weak a possible trigger of word order shift. But if Num and, more importantly, Adj underwent reanalysis – the former becoming a normal Mod, the latter losing its verbal characteristics — the inherent mobility of the Adj certainly facilitated its left-dislocation in analogy with the other light Mod s, Dem and Num.

The strong connection between Adj s and V s in EC (the former being a subset of the latter) can perhaps be held responsible for the much more difficult and rare N Rel > Rel N shift, with Gen being the last to move. Our hypothesis requires a historical stage with Adj N, N Gen word order, as currently witnessed in Bayso. Coupled as it is with postpositions and OV, this pattern runs counter to Hawkins’ (1983) First Implicational Universal which predicts that if a language has SOV word order then, if the adjective precedes the noun, the genitive also precedes the noun.

This is indeed a very rare pattern, although a few cases have been pointed out (e.g., Dryer [1988] reports Djapu [Australian], and another exception is found in the very same Ethiopian area: Tigre). To this list we can now add Bayso (see above, 2.3.). The rarity of this pattern is caused, in Greenbergian terms, by its combination of very dispreferred orders, both recessive and disharmonic: N Gen is disharmonic with Postpositions, and Adj N is recessive against N Adj. Synchronically, we can thus predict that this pattern will be the by-product of a recent or still ongoing word order shift and
co-occur with less dispreferred patterns: both predictions are met with in Bayso.

Being synchronically rare, this pattern will also be diachronically unstable and apt to change. This has indeed happened in the HEC languages and Saho-Afar which have aligned all Mod.s with the postpositional and V-final order, resulting in a maximally harmonic type although both Rel and Adj are in their recessive Mod N order (if a dominant order can be established for Adj at all).

In other words, having shifted both Adj and Rel before the head, the HEC languages and Saho-Afar were in a way compelled to shift the Gen, too. Again, the present evidence from Bayso with its topicalized genitive construction sheds light on what happened in HEC: the Mod was left-dislocated and linked to its head by a pronominal link.

Less spectacular were the developments within LEC. It seems that the verbal character of the category "Adjective" has been lost in a few languages; this is evidenced by the rise of gender/number agreement, obligatory in Arbore and frequent in Dullay. Case agreement, as found in Oromo, is perhaps the end result of this development in which the Adj takes over nominal characteristics and establishes itself as a grammatical category of the language.

This hypothesis predicts that the Cn will be found in N Adj only when the Adj itself has few verbal characteristics, thus patterning more like a nominal Mod than a Rel. This appears to be the case in Arbore where the Adj agrees with the N in number/gender and is preceded by the Cn. In Dullay, on the other hand, which has less developed gender-agreement of the Adj with the N, the presence of the Cn is optional. The Harso-Dobaso dialect (in which no agreement is found) will probably use the Cn before the Adj less frequently than the Gollango dialect in which number/gender-agreement is normal.

Finally, it is predicted that the Num will be Cn-introduced where it has adjectival features (such as gender-agreement with the head) and will, conversely, be found without Cn where it is a noun, thus following the N Gen pattern. This seems confirmed, e.g., by the optional absence of the Cn in Arbore N Num, for which Hayward (1984:211) suggests the possibility of a N Gen analysis. Adoption of adjectival characteristics on the part of the Num is indeed found in Dullay where many Num.s agree in gender with the N (only "I" does in Arbore). It is more problematic to ascribe to this development the general shift of the Num from pre- to postmodifying position in every LEC language except Somali.

The LEC languages have preserved postpositions and SOV basic word order but have now all the Mod.s after the head, thus displaying a maximally harmonic NP word order, although Dem and Num are in their recessive postmodifying position. One can also note that LEC languages conform to a macro-areal tendency of African languages to have all Mod.s after their head irrespective of the Object-Verb and adpositional order (cf. Dryer 1988:189). Although we have restricted our analysis to East Cushitic, "agreeing" Gen.s (in which the Gen shows gender/number agreement with the following N) are the norm also in Agaw languages. According to Hetzron (1980:59), "this is the strongest isogloss connecting Agaw and RVC [HEC]." The author argues that "the use of the k/t elements in genitive constructions in agreement with the possessed seems to be archaic", but he does not give any explanation for this supposed archaicism, nor for its very presence, other than hypothesizing that "these [elements] served as obligatory or quasi-obligatory determiners of the possessed (which is also considered obligatorily definite in the Semitic construct state) [...]. The innovation of Highland Cushitic (Agaw - RVC) consisted of transferring the k/t element, originally a definitizer that had become a regular genitive-marker, to the possessor, while keeping the original agreement solidarity with the possessed" (Hetzron 1980:60f.).

The hypothesis sketched in this paper seems to better account for the rise of this unusual genitive pattern without having recourse to an improbable shift of the definite marker from the possessed (the N) to the possessor (the Gen) for unspecified "typological reasons" (Hetzron 1980:61).

References


Mauro Tosco


Emeneau, M.B. 1956. India as a linguistic area. *Language* 32:3-16.


