ISSN 0165—4004

FOLIA LINGUISTICA

Acta Societatis Linguisticae Europaeae

Tomus XXIII/1–2

1989

SONDERDRUCK

OFFPRINT

TIRAGE A PART

MOUTON/DE GRUYTER

SOCIETAS LINGUISTICA EUROPAEA

Since the elections of 1988 in Freiburg/Br. the following persons now serve as officers or members of Committees:

President:		Mario Alinei, Zeist
Vice-President:		Paul Valentin, Paris
Secretary/Treasurer:		Werner Winter, Kiel
Editors:	FoL:	Wolfgang U. Dressler, Wien
	FoLH:	Jacek Fisiak, Poznań
Executive Committee:		Francisco Marcos Marín, Madrid E. F. Konrad Koerner, Ottawa Harm Pinkster, Amsterdam Richard J. Watts, Bern Fred Karlsson, Helsinki Maria Iliescu, Bucureşti Dieter Kastovsky, Wien
Publications Committee:		Jacek Fisiak, Poznań Martin Harris, Salford Irene Philippaki-Warburton, Reading Wolfgang U. Dressler, Wien Jadranka Gvozdanović, Amsterdam Ferenc Kiefer, Budapest
Nominating Committee:		Martti Nyman, Helsinki Anna Giacalone Ramat, Pavia Martin Harris, Salford Anders Ahlqvist, Gaillimh Alberto Nocentini, Arezzo

Correspondence concerning membership and other administrative matters should be sent to: Prof. Werner Winter
Olshausenstr. 40—60
D-2300 Kiel
West Germany

Fees and payment modalities:

The annaul membership dues are DM 35.00; the admission fee is DM 9.00 Payment should be made not later than 1 March of the year for which membership rights are desired to the following account:
Postal checking account nr. 2428 67-201, Postscheckamt Hamburg (Professor W. Winter, Kiel).

If payment is made in a currency other than German marks, the equivalent of DM 35.00 (or DM 9.00, respectively) should be increased by ten percent to cover conversion charges.

Best wishes,
Robert

TYPOLOGICAL PECULIARITIES IN SOMALI[1]

ROBERT HETZRON

0. INTRODUCTION

Linguistic typology has as its essential object the crosslinguistic study of what external shape some given linguistic phenomena are capable of assuming. As a corollary, it will also be established then what logically possible shapes or manifestations are excluded. Moreover, a compilation of typological facts has often led to the recognition that there may be a correlation between the formal manifestation of a certain category and the formal manifestation of another.[2] In this manner, a deep, often otherwise unsuspected connection between the two categories may be posited, providing thereby a challenge for explaining the rationale for this correspondance. The delimitation of possible vs. impossible formal expressions and the discovery of a deep connection between categories where the relationship does not appear at the first glance may provide important insights into the inner workings of the language.

Naturally, one cannot state with absolute certainty that a potential formal expression cannot exist in reality just because it has not (yet?) been found. For that, one should not only examine all the languages of the world, a technically impossible task (recent work on Hixkaryana, a Carib language of northern Brazil, has just illustrated what may happen: at long last an OVS language was discovered), but one should also show through some concrete means why such a form could have never occurred and will never occur — a chimerical enterprise. Yet, if a thorough study of a sufficient spread and number of geographically, genetically and culturally different languages reveals that a logically possible expression is simply unattested, one may safely state that such an expression, if found somewhere after all, as in Hixkaryana mentioned above, could at least be labelled "unusual" or "anomalous", and that is

Folia Linguistica XXIII/1–2

0165–4004/89/0023–7 $ 2,–
© Mouton/de Gruyter, Berlin
Societas Linguistica Europaea

already an achievement. Even if it does show up in the future, it will have to be treated as something special, in need of special notice or special explanation. The generalizations that had been based on its non-existence are not necessarily invalidated. At worst, they will be demoted from the rank of "Absolute Universals" to the still respectable status of "Strong Tendencies", but even that may not be necessary.

A case in point is found in Awngi (Southern Agaw), a Cushitic language. Greenberg's Universal No. 43. states that "if a language has gender categories in the noun, it has gender categories in the pronoun." (Greenberg 1963 : 96 and 113). Awngi is an exception to this, — it has masculine and feminine in nouns (e. g. *muliqísí* 'monk'/*moleqésá* 'nun'), also in verbs (*desγʷà* 'he studied'/*destəγʷà* 'she studied'), but there is no gender distinction in pronouns, even Sg. 3. has *ŋi* for both 'he' and 'she'. But this is one of the cases where exceptions, against all logic, confirm the rule. This anomalous situation demonstrably (by comparing the various Agaw languages) came about in the following way. Originally for 'he/she' there was **ni*/*nit* with the usual feminine ending *-t*. A well-attested conditioned sound change $t > r$ made the feminine into *nir* or the like (attested in Bilin, Xamir). The non-northern dialects further had a conditioned sound change $r > y$ (perhaps only for *r* that comes from *t*, which would indicate that it was not a simple *r*), which made 'she' into *niy* or *ni*: in Kemant and Quara-Falasha, barely distinguishable from *ni* 'he', especially in languages that have no vocalic length. Therefore, Awngi could not help merging its third person singular pronouns. Through a phonetic accident, the forms of 'he' and 'she' ended up sounding the same. Maybe Greenberg's 43rd Universal is weakened by this to the extent that if it were such a compelling principle, one would expect Awngi to develop some new pronouns to maintain gender-distinction, but this is far from being a fatal blow to Greenberg's intuitively and, with one exception, empirically correct generalization.

Now, we try to figure out the inventory of logically possible expressions of a category by a preliminary examination of the facts, followed by some abstract thinking. We find a number of expressions that constitute a class. Then, through logical considerations, we try to establish a complete set of items that may constitute the full extent of that class. In doing that, we may find that some of the possibilities available are not exploited by natural languages.

Yet, in some cases, we may find certain formal expressions that do not fit into the framework of the class that has been established on the basis of the examination of the frequently attested situations. The introduction of these special types of expressions will indeed have an impact on the general format of the relevant typological statement. An example of this is given in 2.2. below.

In the following, some typologically interesting phenomena of Somali are presented. I shall try to point out some peculiarities of its structure that are rare among the languages of the world, and show, in some cases, that apparent exceptions to empirically established universal laws are not real exceptions, but are secondary as evidence, for they depend on other constructions.

1. THE CASE OF "r"

In Somali, word-initial or geminated *r* is pronounced voiceless; in other positions, i. e. ungeminated and not in the beginning of the word, they sound voiced, of the common "rolled" type. This is exactly the same distribution that we find in Classical Greek for the *spiritus asper*, the accent over *ṙ* which must have marked the same phenomenon as also confirmed by the Latin transcriptions of Greek words, also used in English, initial *rh-* and geminate *-rrh-* (as in *rhythm* and *diarrhea*). The phonetic differentiation of geminate and nongeminate *r* is known from other languages, even in those that otherwise have no more gemination. Such a separation in the articulation may be the origin of the uvularization of French *r*. Spanish has maintained a distinction between *r* and *rr*, but I do not know of other languages where initial and geminate *r*'s are treated the same way.

2. CASE-MARKING

In many languages, case-marking is not a morphologically homogeneous phenomenon. This has led to a great deal of confusion as to what is to be considered a "case". For example, in German one speaks of four cases, because nouns have four kinds of endings, for Nominative, Accusative, Dative and Genitive. From a functional point of view, however, cases are either solely expressed by those

endings, or, more often, by a combination of a preposition and one of the endings. Actually, the domain of the last two suffixal cases is being curtailed by competing preposition + ending combinations, *an* + Accusative for the dative and *von* + Dative for the genitive (cf. Hetzron 1980b). Similar statements may be made of Latin, Slavic, Classical Semitic, etc. The situation in Somali is also not without complications.

2.1. Postpositional case-marking

The position of case-markers with relation to the nouns they situate in the sentence has been one of the pivotal typological criteria. A good correlation has been found between this and the relative position of the verb and the object noun.[3] In casual mentions Somali has received in typological listings (e. g. Greenberg 1963 : 110), it was classified as "postpositional", i. e. having its case-markers f o l l o w i n g the noun.

While such an assertion is not entirely without foundation, it definitely is an unfair representation of the true situation. One phenomenon that may misleadingly impress the casual observer is the word-final case-marking in conjunction with the definite article: masc/fem. *-ku/-tu* for non-focused subject, *-ka/-ta* for the rest, or, for referential articles (approximately meaning 'the above-mentioned'), *-kii/-tii* for non-focused subject, but with High tone *-kíi/-tíi* for the rest, e. g. *nin-ku seyf-ta wuu qaadayaa* 'man-the+Masc.+Subj. sword-the+Fem.+Nonsubj. he he+is+taking' for 'the man is taking the sword'.[4] First of all, note the low informative value of this distinction: marking non-focused subjects only. Secondly, it is the deictic *-k-/-t-* elements that convey the major function of the suffixes. No case-marker *-u/-a* would appear without them. It is these deictic elements that have the case-marker, not the noun. The relevant point in this phenomenon is thus the post-noun positioning of the deictics. The final position of the case-marking within the deictic complex is of more limited typological importance.[5]

Nevertheless, Somali has at least one instance of post-noun case-marking as well (not counting relative clauses). Feminine nouns ending in a consonant have a suffix *-i* for the subject case. Distributionally this is very limited. It is clearly residual, the only survival of an old suffixal case-marking system (Hetzron 1980a,

3.2.2.1.). See also 2.3. for postpositions of no direct typological relevance.

In the first paragraph of 2., I have already summed up the situation in German where one wonders what could be considered the legitimate representative of case-marking in typological characterizations. In the context of the above misrepresentation of Somali, one may evoke the situation in Persian which is usually unhesitatingly taken to have a prepositional type of case-system, even though it has residual suffix *-ra* to express definite accusative. In the same vein, the vestigial endings, post-deictic *-u* vs. unmarked *-a*, good Afroasiatic survivals (Hetzron 1980a, 3.2.1.1.) and post-feminine *-i* (an original gendermarker?) ought not be used as decisive criteria.

2.2. Preverbal case-particles

Where is the bulk of case-marking to be found in Somali? In preverbal particles that are not in direct contact with the noun they refer to. In this system, the accusative and what one would expect to be dative remains unmarked. For the rest, the following preverbal case-particles occur: *u* benefactive-allative, *ku* locative-instrumental, *ka* ablative-detrimental, and *la* comitative.[6] Here are some examples (from Puglielli 1981b):

Cali$_i$ waan u$_i$ imid 'Ali$_i$ I+particle to$_i$ I+went' = 'I went to Ali'

xisaab-ta$_i$ buu ku$_i$ xun$_i$ yahay 'arithmetics$_i$-the Focus+he in$_i$ bad he+is' = 'He is weak in math'

Cali baa geed-ka$_i$ laan ka$_i$ goynayay
'Ali Subj. Focus tree$_i$-the branch from$_i$ he-was-cutting' =
'Ali was cutting off a branch of [from] the tree'

Cali$_i$ baa xamar La la$_i$ tagay
'Ali Focus Mogadishu one(impers.) with he-went' = 'One went with [One took] Ali to[unmarked for case] Mogadishu'

For other case-relations, see 2.3.

This means that what seems to be the main thrust of case-marking in Somali falls neither in the class of prepositions nor of postpositions. It constitutes a type of its own: "preverbal case-marking".

Such preverbal case-marking must have developed in proto-Eastern Cushitic (Hetzron 1980a, 3.4.1.), but it broke down in most languages of this group. Somali and Dasenech (Sasse 1976 : 208) have conserved it the most faithfully.

Associating case-marking with verb is naturally not unique to this group of languages. Indo-European and other verbal prefixes of the *in-*, *ex-*, *con-*, etc. type come to the mind, but these are primarily adverbial modifiers of the verb and not replacements of adnominal case-marking. But case-oriented verb-derivational systems are also attested. The applicative in Bantu, Nahuatl and, less consistently, Arabic, etc. make a dative object into an accusative., e. g.

Arabic	Swahili
kataba ilaɣ-r-raǧuli maktu:ban	*niliandika karua kwa mtu*
'he + wrote to the-man letter-Acc.'	'he + wrote letter to man'
vs.	
ka:taba-r-raǧul-a maktu:ban	*nilimwandikia mtu barua*
'he + wrote + Appl. the-man-Acc. letter-Acc.'	'he + wrote + him + Appl. man letter'

'He wrote a letter to the man ~ He wrote the man a letter'[7]

The closest we can come to this in English is the comparison between the expressions *be friends with* vs. *befriend* + direct object, or *wait for* + Noun vs. *await* + Noun, or even *weep over* + Noun vs. archaic English *beweep* + Noun. In this context one may also mention the "casal" verbs which compete with real case-markers (only the former are used in "serializing" languages), e. g. *use* + direct object which overlaps with the occurrence of the instrumental *with* (though in Semitic, the instrumental case must be used also in conjunction with the verb 'use', Hebrew *hištameš be-*, Amharic *bä- . . . täṭäqqämä*). A wider application of this principle, maybe comparable to the Somali situation, is found in the Arawakan language Machiguenga, where case-markers are incorporated into the verbal word (Gary-Keenan 1977 : 104—5, unfortunately no actual data on this are available to me). In Nilotic languages (Tucker-Bryan 1966 : 409—11, 451—2), e. g. Kipsigis *til* 'Cut!', *til-ji:n* 'Cut for . . . !', *tilén* 'Cut with . . . !.' In Fula (Arnott 1970 : 340—70) the Locative and the Dative-Benefactive are marked on the verb and the Comitative on both the verb and the noun:

ɲo-wolwii 'he spoke', but
*ɲo-wolwa**n**ii moodibboo* 'he spoke to the teacher'

with *-an-* for the Dative, and

ɓe-nastii 'they came in', but
*ɓe-nast**id**ii ʔe Bello* 'they came in with Bello'

where the verbal extension *-id-* and the adnominal preposition *ʔe* conjointly express 'with'. For the Locative-Instrumental (McIntosh 1984 : 74):

ɓe-kaɓɓ-ii-mo 'they-tie-VoiceAspectAffirmative-him' for 'They tied him up', but
ɓe-kaɓɓ-ir-ii-mo ɓoggi 'they-tie-in/with$_i$-Voice-Aspect-Affirmative-him ropes$_i$' for 'They tied him up with ropes'

The essence of these systems is that the modification of the verbal form shows which complement, an instrumental, beneficiary, etc. will be the formally unmarked object. One should also mention the case of the Philippine languages where verbal prefixes show what is the function of the focused, formally unmarked noun (Dik 1980 : 88—97), e. g. Maranao

sombaliʔ o mamaʔ sa karabao
'butcher Agent man Object carabao'
the focusless construction for 'The man butchered the carabao'

with Agent-focus:

somombaliʔ so mamaʔ sa karabao
'butcher + Agent. Der. Focus man Object carabao'

with Object-focus:

sombaliʔ in o mamaʔ so karabao
'butcher + Obj. Der. Agent man Focus carabao'

or Tagalog (Schachter-Otanes 1972: 60 and *passim*):

Sinulat ko ang lihan
'wrote I the letter'
('I wrote the letter') vs.

Sinulat-an ko ang titser
'wrote-Dative I the teacher'
('I wrote to the teacher').

However, to my knowledge, Somali and its close relatives are the only languages which operate with a system of preverbal, thus phonologically more independent, particles, very importantly: more than one of which may occur before the verb.[8] This is not a device conferring special status to one noun (as focus in Philippine languages, unmarked indirect object, but no other case, in English), but a full-fledged case-marking system that can be manipulated as such, using thus a typologically unique method.

2.3. Postpositions in Somali

There are, however, adnominal postpositions in Somali for more complex relations, e. g. (Bell 1953, § 78):

geedkii agtiis-ii 'tree$_{masc}$-the proximity-his-the [= near]'
daarta gudaheed-a 'house$_{fem}$-the inside-her-the [= inside]'
nimankaas hortood-a 'men-those front-their-the [= in front of]'

These postpositions are all genitival nouns exhibiting the expected possessive pronominal agreement with the preceding noun. Therefore their postpositional status may not be truly relevant typologically. It is the consequence of the fact that possessee nouns stand after the possessor in Somali. It is significant, however, to remember that of the several possible genitival constructions, it is the one that allows postpositional status that has been picked for these adverbial expressions.

Like many other languages, including English, Somali has two basic genitival constructions on the basis of the order of constituents (Bell 1953, § 77), and the Possessor-Possessee construction may be further subdivided according to the criterion of definiteness of constituents (Caney 1984: 312—3 and in greater detail Gebert 1981: 55sq). Here are the examples. They are mainly taken from Gebert 1981, the distinction between (i) and (j) was suggested by Hyman (1981: 129), (q) (r) and (s) come from Caney 1984: 58—9 and *passim*:

(a) *macallin guri-giis* 'teacher house-his' ('a teacher's house')
(b) *macallin-ka guri-giis-a* 'teacher-the house-his-the' ('the teacher's house')
(c) *guri macallin-ka* 'house teacher-the' (= (a))
(d) *guri-ga macallin-ka* 'house-the teacher-the' (= (b))
(e) *guri macallín* 'house teacher' ('a house for teachers')
(f) *guri-ga macallín* 'house-the teacher(-Genitive)' ('the house for teachers')

(g) *guri-ga naag-ta* 'house-the woman-the' ('the woman's house')
(h) *guri-ga-naag-oód* 'teacher-the women-Genitive' ('a house for women')
(i) *šimbir bad-eéd* 'bird sea-Genitive' ('[a specimen of] a seabird')
(j) *šimbir bad-éed* 'bird sea-Genitive ('a seabird [generically]')
(k) *šimbir-ta bad-eéd* 'bird-the $_{\text{fem.}}$ sea-Genitive' ('the [specific] seabird')
(l) *šimbir bad-éed-da* 'bird sea-Gen.-the $_{\text{fem.}}$' ('the seabird [as a category]')
(m) *kalluun-ka bad-eéd* 'fish-the$_{\text{masc.}}$ sea-Gen.' ('the seafish [spec.]')
(n) *kalluun bad-éed-ka* 'fish sea-Gen.-the$_{\text{masc.}}$' ('the seafish [generic]')
(o) *bixid(-da)dhiig* 'exit(-the)blood 'hemorrhage')
(p) *diig bixid(-da)* 'blood exit(-the)' (= (o))

and to be complete

(q) *dhar dagaal* '(clothes war' ('uniform')
(r) *Ciidan-ka bad-da* 'force-the$_{\text{masc.}}$ sea-the$_{\text{fem.}}$' ('navy')
(s) *Ciidan-ka difaaC-a cir-ka* 'force-the$_{\text{masc.}}$ defense-the$_{\text{masc.}}$ sky-the$_{\text{masc.}}$ ('air defense force')

The following analysis is a summary of Gebert's conclusions, completed with Hyman's observations. Examples (a) and (b) contain the Possessor-Possessee order where the latter is followed by a possessive pronoun coreferential with the possessor. In (a) the Possessor is indefinite, whereas the possessive pronoun automatically makes the Possessee definite: 'A house of the teacher'. In (b) both nouns carry an article, the second one by agreement only, the meaning is 'the teacher's house', i. e. 'the house of the teacher'. All the other examples are of the Possessee-Possessor order. (c) is the perfect equivalent of (a), 'a house of the teacher', whereas (d) corresponds to (b) 'the house of the teacher'. The forms (e) and (f) belong to another construction, exemplified by (i) and (k) (see note 10 below). (g) is like (d), with both nouns definite, where the Possessee could also drop the article, but (h), the feminine, has the genitive ending. The genitive ending in -VV*d* appears after feminine nouns, collective plurals with different vowels, masculine nouns only have a tonal modification to mark the genitive, a high tone appears in the last vowel of a genitival noun which, in an indefinite form, follows the Possessee-Headnoun ('teacher' would be *macállin* otherwise).[9] Constructions with genitive-ending are to be understood as compounds where the Possessor is a qualifier with no specificity of its own. Here, however, a further subdivision was discovered by Gebert (1981: 67—9) for genitival constructions containing an article (k — n), and with even more refinement, taking tone also into account, by Hyman (1981: 129—30). The rule that places a High tone on the last syllable of the genitival noun is valid here as well, so that (i), (k) and (m) will have it as well. This will be a general

possessive construction where the possessor one is a qualifier, with no identity of its own, only serving to limit the domain of the entity marked by the headnoun to those so qualified, while the headnoun will still refer to a specific entity. In this case, the definiteness of the headnoun will be marked on the headnoun itself, (i) vs. (k) and (m). On the other hand, by not having the final tone (but a penultimate one) another construction arises that deprives even the Possessee-Headnoun of its specificity, and the whole expression becomes generic: (j), a seabird as a category of birds or animals or denizens of the sea area. In this latter case, the article will follow the entire expression: (l) and will actually agree in gender not with the genitival noun preceding it, but with the more distant headnoun: (n) where 'fish' is masculine, 'sea' is feminine, while the article following the latter is also masculine. Hyman considers these true compounds, basically almost lexicalized single nouns.[10]

Examples (o) and (p) are not essential for our discussion. I list them all the same because they are interesting. (o) is a Possessee-Possessor order of the (e-f) type, with the definite article allowed. (p) is the exact reverse, with the exceptional order Possessor-Possessee, with the article still after the Possessor. The whole expression means 'hemorrhage'. As pointed out by Gebert (p. 73), the (p) pattern is possible only with the deverbal nouns and it is a remnant of the finite sentence order from which it originates, where the verb is postnominal. Thus *bixid* is not a nominalization of the verb 'go out' by itself, but the entire phrase 'blood go-out' was nominalized as a whole, whereas (o) is a nominalization of the verb *bax-* 'exit', placed, afterwards, in a compound-genitival relation with 'blood'.

The last three examples are from Caney (1984), though Gebert (1981) has some of them, but without assigning them a separate status. They give me the impression of being a separate construction where either all (p-q) or none (o) of the components carry the article. If they do, it appears always according to the gender of the marked element (e. g. the genders differ in (r), but all the set of articles counts as a single definitization). I could not find out whether the phrasal groups marked indefinite (on all their components) may be definitized at all, and if not, what determines the definiteness vs. indefiniteness of each group. But the whole process of grouping seems to be an innovation, and does not properly belong to our discussion.[11]

We may then summarize the genitival constructions available to Somali:

(A) POSSESSOR$_{i.indef/\langle def\text{-}j\rangle}$-POSSESSEE + Poss.pr.$_{i\langle +\text{-}a_j\rangle}$[12]
(exx. (a-b))
(Aiii) POSSESSOR$_{indef}$-POSSESSEE$_{(deverbal)indef/def}$
(ex. (p) = (Biii))

(B) POSSESSEE-POSSESSOR orders
(Bi) POSSESSEE$_{indef/def}$-POSSESSOR$_{def(!)}$
(exx. (c-d) and (g) = (A))
(Bii) POSSESSEE$_{indef/def}$-POSSESSOR-Genitive
(exx. (e-f), (h-i), (k), (l) (o))[13]
(Biii) [POSSESSEE-POSSESSOR-Genitive]$_{indef\text{-}def}$[14]
(Biv) POSSESSEE$_{(deverbal)indef/def}$-POSSESSOR$_{indef}$
(ex. (m))
(Bv) POSSESSEE$_{\langle indef/def\rangle i}$-POSSESSOR$_{\langle indef\text{-}def\rangle i}$
(exx. (o-p))[15]

Correlations between definiteness and alternative possessive (genitival) constructions are known from many languages. In Semitic the "construct state" automatically makes the possessee definite. It is clear that the above genitival postpositional construction h a d to select construction (A) to allow the case-marked noun to appear either definite or indefinite, (the definiteness of the postposition-possessee is vacuous). Therefore, the postpositional character of these elements is of no direct typological relevance, it only follows from the necessity of adopting construction (A). As we shall see in Section 4, (A) is a more fitting order for a verb-final language, but I believe that this was not the major factor in opting for it, but rather the freedom of assigning either value of definiteness to the noun was the decisive element.

3. THE ORDER OF CASE-MARKING

3.1. The order of case-marking particles

There are four case-marking preverbal particles in Somali. Three of them are multifunctional. For example, *ku* may mark the instrumental as well as the locative, and in principle the two may

cooccur in the same sentence. Moreover, such particles may also appear as part of the lexical representation of a verb. Thus, 'to return' is not merely *noqo* in Somali, but *ku noqo*, with a *ku* carrying by itself no special meaning. Therefore, it is not surprising to find that three out of the four preverbal particles may occur twice in the same sentence. The major limitation is that no more than three such particles may be used in one sentence. The following is a representation of the relative order of case-markers, where 2 indicates that that particle may occur twice (Hetzron 1977b: 200):

U^2-*KU*^{2}-*KA*^{2}-*LA* 'Benefactive $\sim$ Allative-Locative $\sim$ Instrumental-Ablative $\sim$ Detrimental-Comitative'

out of which any three may be selected, in this order.

It should be noted that the cooccurrence of several particles will have morphophonemic consequences. Non-initial *k* becomes *g*, non-initial *u* is represented by *gu*, and *ku*, when combined with another *k*-containing particle (another *ku* or *ka*) will adopt a vowel *a* (*kaga* for *KU*+*KU* and *KU*+*KA*, in addition to *KA*+*GA*, thus three-way ambiguity arises). Further complications appear when complement pronouns are prefixed to the case-particle complex. For a detailed treatment, see Andrzejewski 1960. The implications of this phenomenon for a general theory of clitic particles have been treated in Hetzron 1977b: 199—200.[16]

3.2. The order of case-marked nouns

Since case-markers are not contiguous with the nouns they refer to, the question arises how it is indicated, in the event that more than one preverbal case-marker is used, which case-marker belongs to which noun. The answer to this question is far from simple, for various reasons. The easiest solution is not to stack up too many nouns in one sentence, so that no confusion may result. Somali is not a true serializing language, but instead of having a long line of complements depending on one verb, it prefers having more verbs so that each should support less nouns. Yet, if one tries to force the hand of a native speaker and have him/her produce a longer line of complements, one encounters several difficulties. One of them is

that Somali has become a national language with a general, supra-dialectal koiné being formed. The rule to be given below seems to be true of northern dialects only, but the southern capital city Mogadishu where much of cultural life converges attracts speakers from everywhere, and the younger generation of northerners seems to have given up this delicate rule. Even though years ago I had been in contact with speakers having this feature in their dialect, I have not been able to encounter one since I started to work on this paper, so that my newly arising inquiries could not be answered. Moreover, Somali is a non-rigid SOV language where complement nouns may easily appear after the verb as well, making the rule given below somewhat vacuous. Last but not least, the rule does not work when there is any focus in the sentence, a frequent event in Somali discourse, even in the northern variety of Somali (that I am not able to delimit with any claim for precision). Here is the rule alluded to:

When the sentence contains no focus and all the nouns are pre-verbal, in the northern dialect represented by my friend Ahmed Dirie, other Somali friends, and the informants for the written documents to be quoted below, the major linearization principle consists of a m i r r o r - i m a g e. After the subject and object nouns, complement nouns are aligned in an order opposite of that of the case-marking particle, (i.e. Comitative-Ablative-Locative/Instrumental-Benefactive, see Herzog 1960), schematically:

$\ldots N_i\text{-}N_j \ldots case_j\text{-}case_i\text{-verb}$
and never
$*\ldots N_i\text{-}N_j \ldots case_i\text{-}case_j\text{-verb}$

Here are some examples:

Ninku boqorka af soomaali wuu kula hadlaya (Herzog 1960)
'man-the-Subj. chieftain-the-Oblique$_i$ [language Somali]$_j$ he in$_j$+with$_i$ he+is+speaking'
'The man speaks with the chieftain in Somali'[17]

Ceelka adhig La gaaga soo saarey (Bell 1953, § 115)
'Well-the$_i$ rope$_j$ one you+with$_j$+from$_i$ pulled out'
'They pulled you out of the well with a rope'

Joonisbaag lug kaga baxay (Andrzejewski 1960)
'Johannesburg$_i$ foot$_j$ on$_j$+from$_i$ he+went+out'
'He left Johannesburg on foot'

In the last example, *kaga* stands for *ku+ka*. Our principle may be confirmed by the irreversible order of the nouns: **lug Joonisbaag*.

As mentioned above, the mirror-image principle does not work when focus is involved:

Albaab-k-uu ul cali kaga garaacey (Bell 1053, § 114)
'Door-the-Focus+he stick$_i$ Ali$_j$ with$_i$-against$_j$ he+knocked'
'He knocked at[= Acc.] the *door* with a stick, against Ali's wishes'

Or the structure may be coincidentally mirror-image-like:

Ninkii baan ceelka xadhig kaga soo saarey
'man-the$_{referential}$ Focus+I well-the+Oblique$_i$ rope$_j$ with$_j$-from$_i$ I+pulled'
'I pulled the *man* out of the well with a rope'

I have found one case where *u* doubles, and is reduplicated, as Allative and Benefactive, in either order of the nouns, but the direct object is inserted between them:

Sarkaalka warqaddan Cali ugu gee (Bell 1953, § 116)
'Officer-the+Oblique letter-this Ali for-to ~ to-for take!'
'Take this letter to the officer for Ali/to Ali for the officer!'

If one may base conclusions on a unique example this far found, when the same case-particle is reduplicated, there is no fixed order of nouns based on function, both readings are correct ('for' means, generally, 'for the sake of', 'to the benefit of', whatever the pragmatic implications).

3.3. Complement pronouns

When complements are represented by unstressed pronouns, these are prefixed to the chain of the case-marking particles, e. g. *i-la* 'me-with' for 'with me'. With two pronominal complements, pronouns homophonous with the masculine possessive endings are inserted between the case-markers and the verb: *ku-ga key* 'thee-

from mine' = *i-ga kaa* 'me-from thine', both for either 'me from you' or 'you from me'. Thus, it is unclear which one of the pronouns stands for the object and which belongs to the case-marker. In this respect the following regularities have been this far observed (at least in the speech of the informant for Herzog 1960).

In combination with the benefactive *u*, the prefixed pronoun may either refer to the object or the complement: *ii* (for *i* + *u*) may either be 'for me' or 'me for NOUN/him/her'. The same is true of the ablative *ka: i-ga* 'from me' or 'me from NOUN/him/her'. On the other hand, in combination with the locative-instrumental *ku*, the pronoun may only refer to the accusative object: *igu* 'me in~on/ with NOUN/him/her' and never *'in/with me'. In contrast to this, when combined with the comitative *la*, the pronoun always refers to the complement: *ila* 'with me' and never *'me with NOUN'. The restriction on the latter two seem to be of a pragmatic nature. Instrumentals and locatives are unlikely to have human reference, which is the essence of the pronouns,[18] whereas comitatives are practically always human.

4. THE POSITION OF ADJECTIVES

In Somali, adjectives follow the noun they qualify. Greenberg (1963: 79 § 85) only makes relatively weak statements on this issue: "If a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun" (Universal 5), and summarizes the situation in 11 SOV languages examined: Noun-Adjective order in 5 of them, Adjective-Noun in 6 (Table 5). Now, Somali has both orders of genitival constructions (see 2.3. above), so that Universal 5 is neither contradicted nor confirmed. It is even difficult to decide which one of the genitive constructions should be considered unmarked. However, intuitively one feels that just like construction (a), the Possessor-Possessee linearization, the Adjective-Noun order is more natural for a verb-final language and it is the other one that requires explanation (see Vennemann 1974: 345, following the principle, labelled "the principle of natural serialization" which says that the relative order of the "Operator" and "Operand" should be consistent throughout the grammar of a language (*ibid.* 347)). How did then the not uncommon, yet less natural Noun-Adjective order come about?

The same configuration, verb-final sentence order and post-noun adjective is found also in Yuman languages of the South-Western United States and has been studied by Langdon (1977). She showed that these adjectives are of verbal origin and they act as predicates even synchronically, hence their position.

For most adjectival qualifications, similar logic may be applied in Somali as well. Somali does have its relative clauses in a post-nominal position. One should also remember that most adjectives in Somali do indeed appear in relative clauses only, they can only be predicative and when their function would be to qualify a noun, they have to be accompanied by a short copula. Bell (1953) distinguishes several types of adjective. "Radical adjectives", a small number of qualifiers (he lists 16 of them), may stand alone after the noun (§ 83). He suggests that these are "nouns in the first place, and their use as adjectives appears to be no more than an extension of the possessive use", p. 76, note 1). Next he mentions "verbal adjectives" which "can be formed from any verb, if the sense makes it possible" (§ 84), then, the most important class "compound adjectives" formed of a noun and *ah* 'who is', *leh* 'who has' and *la?* 'who does not have' (§ 85). Andrzejewski (1969) uses the term "hybrid verbs" for these. Some examples: *bir ah* 'made of iron' (*bir* 'iron'), *roob leh* 'rainy' (*roob* 'rain'), *lug la?* 'legless' (*lug* 'leg'). It does seem then that the most frequent type of noun-qualifying is a relative clause which is naturally placed after the noun. We may add here that, as illustrated in the genitive construction (p) in 2.3. with a deverbal Possessee headnoun in final position just as a trace of its verbal origin and verbal position, Somali does seem to exhibit a great deal of "loyalty" to the origin of its constructions. On the other hand, one may hardly claim that the "radical adjectives" are also underlying relatives. One cannot accept Bell's interpretation of them as possessor nouns because when they carry no article on their own, they should be, if they were nouns, followed by a genitive ending, as in pattern (Bii/iii) in 2.3. Yet this is not the case: *sananadda dhow* 'year-the$_{fem.}$ near', for 'the next year'. Therefore, all we could say is that the few true adjectives merely follow the standards of the truly frequent predicative class of compound or derived adjectives. This would then be the result of the tendency for uniformity.

5. THE POSITION OF THE NUMERALS

As already pointed out in different contexts, modifiers are expected to precede the noun in verb-final languages. While the Somali adjectives do not abide by this principle, numerals do, — they precede the noun quantified. This is again a conflict situation. Greenberg (1978: 77) mentions the "almost exceptionless word order implication N[oun] Q[uantifier] → N[oun] A[djective]". Now, he adds "We also assume that a language will not simultaneously change both noun—quantifier and noun—adjective order." Therefore, a QN. NA order may arise from either QN . AN or NQ. NA (*ibid.*, Figure 9). Greenberg further wonders, which one of the two terms is "the more dynamic factor" and speaks of the "putatively greater independence of the adjective-noun order" (*ibid.* 78) which would lead to the assumption that the position of the quantifier represents the more archaic situation.

Actually, the positioning of numerals is also determined by another factor, that of the genitive. In construction (h), etc. of 2.3., where the genitive ending -VV*d* when it is a plural: *dhal shimbir-eed* 'a fledgling' ('a descendant of a bird'), *hadal naag-ood* 'talk of women'. With numerals, feminine nouns not ending in *-o* also have an ending *-od* added: *lix habl-ood* 'six girls-Gen.' (Bell 1958, § 51). This suggests that the numeral construction, whether this is synchronically true or not, started out as a genitive structure: 'six of girls'. This happens to be exactly the case in the VSO-type Classical Semitic: adjectives and possessors follow the noun whereas numerals precede it, but numerals also require, at least in the archaic portion of numeration (thus excluding the innovative "-teen" numerals of Arabic which require the accusative, see Hetzron 1977a), the genitive case. Thus, adjectives and numerals do not have to go together, whereas numerals and genitives do. The interesting point is that, as in the case of the postpositions (2.3.), the Head—Modifier (Operand—Operator) relationship is reversed, the Possessor noun becomes the Head, the Possessee the Modifier. This means that the relative order of the numeral and the noun is also likely to change, because of the reinterpretation of the Head—Modifier relationship makes these, according to Vennemann's principle of natural serialization, end up on the wrong side. And this is precisely what happened in the closely related Rendille which did make an adjustment and uses now numerals in a post-noun position, e. g. *albé liḥ* 'girls six' for 'six girls'.

6. SOME FINAL WORDS

The aim of this contribution was to show the typological significance of some facts of Somali grammar, what they can mean to the general linguist. Somali has a number of unusually interesting phenomena, part of which have been summed up above.

Address of the author: Robert Hetzron
Dept. of Germanic
University of California
Santa Barbara
California 93106
USA

NOTES

[1] An earlier version of this paper was originally presented on 21 April 1980, at the Colloquium on Language Typology organized by the Koninklijke Nederlandse Akademie van Wetenschappen, in Amsterdam. The present version is an updated one, with a considerable amount of changes and corrections. Observations on the manuscript by Professors C. Douglas Johnson and Sandra A. Thompson helped to improve it, I express my gratitude to them.

For the Somali examples I am using the official orthography, the only true weakness of which is not marking tones. I shall use tone-marks only when specially necessary. Otherwise the conversion to phonetic syllables is the following: *c* = the laryngeal (ʿ), *dh* = the retroflex (ḍ), *sh* = (š), *x* = the laryngeal (ḥ). Long vowels are marked doubled and capitalization is used prescribed by orthographic conventions. The only deviation will be the marking of morpheme-boundaries by means of a hyphen "-", parallelled by "-" in the morpheme-translation where " + " connects the translations of the components of a *portmanteau* morpheme unseparated in the original, or a set of morphemes the separation of which is not relevant for our topic.

[2] I.e. if category A has the form X, it is certain or likely, whatever the case, that category B will have the shape Y. This concept of "Implicational Universals" has been associated with the name of Joseph H. Greenberg, launched in his revolutionary 1963 article. Vennemann's and Bartsch's "natural serialization principle", to be dealt with later, is a generalization of word order rule principles on different levels of the language.

[3] First stated, also involving the position of the subject, by Greenberg (1963). For the theoretical implications of this, i.e. the predicate-like character of the case-marker, see Hetzron 1974: 55, Vennemann 1974, etc.

[4] Actually, the subject-focus particle *bàa* and the preceding article **-ka* are contracted to *-kàa*, but the Nonsubject-marking of the focused subject may be proved by observing the referential article: *nín-kii líbàaḥ bùu dílay* 'man the$_{\text{referential}}$+Masc.+Subj lion Focus+he he+killed' for 'The man killed the *lion*', but *nín-kii bàa líbàaḥ dílay* 'man-the$_{\text{referential}}$+Masc.+Nonsubject Subj. Focus lion he+killed', for 'The *man* killed the lion'. This may be a trace of an older cleft construction, 'It is the man who. . .' where "man" is a predicate.

[5] In Hungarian, demonstrative pronouns combined with definite articles are the only elements that agree in case with the noun they are preposed to. In German, definite articles take over the thrust of case-marking, away from the nouns which otherwise carry them. This shows the idiosyncratic status of definitizers/deictics in attracting morphological material.

[6] See 3.1. below for the coocurrence of more than one of these particles in the same sentence.

[7] English also has two versions, but the verbal form is the same.

[8] Fula (Arnott 1970: 367) may have a combination of the Locative and the Comitative, in this order, in the verbal extension system, which also includes categories other than case-markers. Their relative order is determined by their phonetic shape, by the major consonant: T—D—N—R (*ibid.*, 366).

[9] See below for the final tone also in -VV*d* endings. In the Benadir dialect that has no tones, -VV*d* may be added to some classes of masculine nouns as well, e.g. *kab nin* = *kab nim-eed* 'shoe man(-Gen.)' for 'a man's shoe'.

[10] It is not clear whether (e) has its quasi-lexicalized counterpart when no final High tone appears. . .

[11] There may be elements other than nouns involved in such phrasal groupings, and those will receive no article, e.g. *dalal-ka soo kordhaya* 'countries-the hither who+are+increasing' ('the developing countries'), a noun combined with a relative clause.

[12] Here I intend to symbolize by "< >" that while the choice between indefiniteness and definiteness is allowed for the Possessor, the Possessee, automatically definite because of the possessive pronoun, will carry an external mark of definiteness -*a* only through agreement, and the indefiniteness ~ definiteness values will apply to the whole compound expression once.

[13] This is used for compound expressions where the Possessors are reduced to the role of nonspecific qualifiers, yet not to be confused with amalgamated compounds used as neologisms where actually phonologically single words are created, e.g. *magaalomadax* 'capital city' from *magaalo* 'head' and *madax* 'head', (cf. Caney 1984: 224 and passim). This expression is characterized by final High tone on the genitival word.

[14] These are compounds for generic of the whole expression, no final High tone is used.

[15] It seems that in this form, a device called "Phrase-grouping" by Caney (1984) for forming neologisms, either all the components or none carry the article (as symbolized by "$\langle\ \rangle_i$"), which will count as single definiteness-marking for the whole expression that refers to a single entity.

[16] The same clitic may be repeated if the functions assumed are different. Now, since on the surface we may have *ugula*, *ugaga* and *kagala* because of the limitation of no more than three such particles in a row, a strictly linear representation of the surface order of the clitics is impossible, because *ku* and *ka-ka* may not be situated with relation to each other, *ku* and *ka* may never be neighbors, see Hetzron 1977b for details on this and related matters.

[17] According to my original informant from the North, the order of the nouns is irreversible here. A young informant from the Northwest, living in Mogadishu, accepts either order here. This is a testimony of the decadence of the rule in Somali, the national language.

[18] Note that the third person complement pronouns are zero in Somali.

REFERENCES

Andrzejewski, B. W. 1960. "Pronominal and prepositional particles in Northern Somali." African language studies 1, 96—108.

Andrzejewski, B. W. 1969. "Some observations on hybrid verbs in Somali." African language studies 10, 47—89.

Arnott, D. W. 1970. The nominal and verbal systems of Fula. Oxford: Clarendon Press.

Bell, C. R. V. 1953. The Somali language. London: Longmans.

Caney, John Charles. 1984. The modernisation of the Somali vocabulary with particular reference to the period from 1972 to the present. Hamburg: Buske.

Dik, Simon. 1978. Functional grammar. Amsterdam: North Holland.

Gary, Judith Olmstead and Edward Louis Keenan. 1977. "On collapsing grammatical relations in universal grammar". In: Syntax and Semantics. 8. Grammatical relations. New York: Academic Press, 83—120.
Gebert, Lucyna. 1981. "Il sintagma nominale", In: Puglielli 1981a, 45—137.
Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements", In: Greenberg, Joseph H. (ed.) Universals of language. Cambridge: MIT Press, 73—113.
Greenberg, Joseph H. 1978. "Diachrony and language universals", In: Joseph H. Greenberg, (ed.) Universals of human language. Stanford: Stanford University Press. Vol. 1. 69—91.
Herzog, Robert. 1960. "Les compléments verbaux en Somali", Comptes Rendus du Groupe Linguistique d'Études Chamito-Sémitiques 8, 92—95.
Hetzron, Robert. 1974. "A synthetic-generative approach to language", Linguistics 138, 29—62.
Hetzron, Robert. 1977a. "Innovations in the Semitic numeral system", Journal of Semitic Studies 22, 167—222.
Hetzron, Robert. 1977b. "Clitic pronouns and their linear representation", Forum Linguisticum 2, 189—215.
Hetzron, Robert. 1980a. "The limits of Cushitic", Sprache und Geschichte in Afrika 2, 7—126.
Hetzron, Robert. 1980b. "On word order and morpheme order". In: Brettschneider, G. (ed.) Wege zur Universalienforschung, Tübingen: Narr, 175—179.
Hyman, Larry M. 1981. "L'accento tonale in somalo", In: Cardona, Giorgio R. and Francesco Agostini (eds) Studi Somali. 1. Rome: Ministero degli Affari Esteri, Comitato Linguistico per l'Università Nazionale Somala, 109—139.
Langdon, Margaret. 1977. "Syntactic change and SOV structure: the Yuman case", In: Li, Charles N. (ed.) Mechanisms of Syntactic Change, Austin: Univ. of Texas Press.
McIntosh, Mary. 1984. Fulfulde syntax and verbal morphology. Boston: KPI.
Puglielli, Annarita, ed. 1981a. Studi somali 2. Sintassi della lingua somala. Rome: Ministero degli Affari Esteri, Comitato Tecnico Linguistico per l'Università Nazionale Somala.
Puglielli, Annarita. 1981b. "Frase dichiarativa", In: Puglielli 1981a, 3—44.
Sasse, Hans Jürgen. 1976. "Dasenech", In: Bender, M. L. (ed.) The Non-Semitic languages of Ethiopia. East Lansing: Michigan State University.
Tucker, A. N. and M. A. Bryan. 1966. Linguistic Analyses. The Non-Bantu languages of North-Eastern Africa. London: Oxford University Press.
Vennemann, Theo. 1974. "Analogy in generative grammar. The origin of word order", In: Heilman, Luigi (ed.) PICL 11, Vol. II, 79—83.

CONTENTS

(continued)

H. H. Chipman and M. Imperiali di Francavilla, His or Hers? How Young English and Algerian Arabic-Speaking Children Understand Third Person Possessive Pronouns in Object Position 157

Cs. Pléh, F. Ackerman and A. Komlósy, On the Psycholinguistics of Preverbal Modifiers in Hungarian: Adult Intuitions and Children's Treatment of Modifiers 181

Ch. Schaner-Wolles, Erstspracherwerb im Sprachvergleich. Zu einem Buch von Dan I. Slobin (ed.) 215

Libri 227

Forthcoming conferences, calls for papers 245

FOLIA LINGUISTICA

Correspondence concerning *Folia Linguistica* should be addressed to the editor:

Prof. Dr. Wolfgang U. Dressler,
Institut für Sprachwissenschaft
Liechtensteinstraße 46a,
A-1090 Wien, Austria

for matters pertaining to *Folia Linguistica Historica* the Editor,
Prof. Jacek Fisiak, Institute of English, Adam Mickiewicz University,
PL-61-874 Poznań, Poland, should be contacted.

The STYLE SHEET for contributions to Folia Linguistica is enclosed periodically with the journal, it is also available from the editor.

For fees and payment modalities, see inside front cover.

PLEASE WRITE NAMES AND ADDRESSES IN BLOCK CAPITALS

CONTENTS

T. V. Gamkrelidze, The "Unconscious" and the Problem of Isomorphism between the Genetic Code and Semiotic Systems (Presidential Address) 1

R. Hetzron, Typological Peculiarities in Somali 7

N. Leko, Case of Noun Phrases in Serbo-Croatian and Case Theory 27

P. Modini, Segmental and Suprasegmental Expressions of Theme and Rheme 55

W. Bublitz, Ausdrücke des Kenntnisnehmens (Hörersignale) oder des Stellungnehmens (Redebeiträge): *yes*, und verwandte Formen 67

SPECIAL TOPIC:
FIRST LANGUAGE ACQUISITION
Guest editor: Csaba Pléh

Cs. Pléh, Foreword 105

J. Pačesová, Language Development: Form and Function .. 107

Ch. Schaner-Wolles, Strategies in Acquiring Grammatical Relations in German: Word Order or Case Marking? 131

(continued inside back cover)

The appearance of the code at the bottom of the first page of an article in this journal indicates the copyright owner's consent that copies of the article may be made for personal or internal use, or for the personal or internal use of specific clients. This consent is given on the condition, however, that the copier pay the stated per-copy fee through the copyright clearance center, for copying beyond that permitted by Sections 107 0. 108 of the US. Copyright Law. This consent does not extend to other kinds of copying, such as copying for general distribution, for advertising or promotional purposes, for creating new collective works, or resale.

© 1989 Mouton/de Gruyter, Berlin; Societas Linguistica Europaea

Printed in Hungary