Studies in Generative Grammar

The goal of this series is to publish those texts that are representative of recent advances in the theory of formal grammar. Too many studies do not reach the public they deserve because of the depth and detail that make them unsuitable for publication in article form.
We hope that the present series will make these studies available to a wider audience than has hitherto been possible.

Editors:
Jan Koster
Henk van Riemsdijk

Other books, still available in this series:

Joan Mascaró and Marina Nespor (eds.)

Grammar in Progress

Glow Essays for Henk van Riemsdijk

CNR inv. 18

Readjustment rules in Somali plural formation

A. Puglielli, *University of Rome*
S. Scalise *University of Venice*)

In a model of Lexical Morphology, Word Formation Rules (WFR's) apply to the words listed in the lexicon, then Inflection Rules (IR's) apply and finally a set of Readjustment Rules (RR's) determines the final shape of the output of the morphological component.[1]

In this paper we will examine the RR's that apply to the output of IR's (namely the Plural Formation Rule (PFR)) in Somali.[2] We would like to show how the operations carried out by RR's crucially depend on three kinds of information: gender, syllabic structure, and phonological structure of the input words.

The first distinction to be made is between masculine and feminine words: as will become clear later masculine and feminine nouns make their plural in different ways, but what is basically different is not the PFR but the RR's that are triggered.

Let us start with the masculine nouns. Within this group, we must distinguish between monosyllabic and polysyllabic nouns and between nouns ending in a consonant and nouns ending in a vowel. Some examples of monosyllabic masculine nouns ending in a consonant are given in (1) (at the right of the arrow there is the plural form):

(1) búug → buugág 'book'
     kóob → koobáb 'cup'
     míis → miisás 'table'
     war   → warár   'talk'

Here, the plural form is made by adding the vowel *á* followed by a consonant which is identical to the final consonant of the singular form. There are at least two ways of describing these data. The first is to suppose that the plural is formed in one step by a PFR in the following way:

(2) $[\_\_\_\_C\alpha]_N \rightarrow [\ [\_\_\_\_C\alpha]_N + áC\alpha]_N$

Condition: N must be monosyllabic

An alternative approach would be to suppose that the PFR only adds the vowel á (cf. 3i) and that, in a second step, a RR 'copies' the final consonant of the input word (cf. 3ii):

(3)   i) $[\underline{\quad}C]_N \rightarrow [\ [\underline{\quad}C]_N + á\ ]_N$
           $\quad\quad$ <+m> $\quad\quad\quad$ <+m> $\quad$ <+m>

Condition: N must be monosyllabic

      ii) $\emptyset \rightarrow C\alpha\ /\ C\alpha á\underline{\quad}\#$

The two possible derivations would then be as follows:

(4)   LEXICON   i)   buug   ii)   buug
      PFR   buug+ág   buug+á
      RR   -------   buug+ág

Among these two solutions, at first glance (4i) seems simpler and indeed it is simpler, but only if we consider the data in (1). In a more general framework, however, (4ii) is preferable because it captures the following generalization: plural formation in Somali consists of the attachment of a vowel plus a RR. This will be made clearer below.

As far as the stress pattern is concerned, we find the following situation:

(5)   búug   →   búugág
      wár   →   wárár

We will thus suppose that the inflectional affix á is always stressed and that a Clash Avoidance Rule (CAR) will operate in order to give the right stress pattern:

(6)   óó → oó

Returning to the plural formation, let us consider now the masculine polysyllabic nouns. These words make their plural in three different ways depending on the phonological structure of the input word. We can say that the affix of the plural for monosyllabic words is always ó (stressed), but then different RR's apply. Consider the following three sets of data:

(7)   i)   doofáar   →   doofaarró   'wild bear'
            sánan   →   sanannó   'superstition'
            sácab   →   sacabbó   'palm (of a hand)'

      ii)   libáax   →   libaaxyó   'lion'
            wáraf   →   warafyó   'catapult'
            kúrus   →   kurusyó   '(camel's) hump'

      iii) máalin   →   maalmó   'day'
            gárab   →   garbó   'shoulder'
            hílib   →   hilbó   'meat'

The plural of the words in (7i) is different from the plural seen in (1) for two reasons: a) ó is added instead of á, b) the 'copied' consonant is 'before' the vowel and not 'after'. Adopting the solution in (3ii) the derivation of the data in (6i) will be as follows:

(8)   LEXICON   sacáb
      PFR   sacáb+ó
      RR   sacáb+b+ó
      CAR   sacab+b+ó

In this case, the rules needed are the following two:

(9)   i) $[\underline{\quad}C]_N \rightarrow [\ [\underline{\quad}C]_N + ó\ ]_N$
           $\quad\quad$ <+m> $\quad\quad\quad$ <+m> $\quad$ <-m>

      ii) $\emptyset \rightarrow C\alpha\ /\ \underline{\quad}C\alpha ó\#$

      where Cα is a stop, a liquid or a nasal

Also the words in (7ii) form the plural by adding ó. What is different, though, is that the RR is not a 'copying' rule but an 'insertion' rule. Of the two rules needed in this case, the PFR (cf. 10i) is the same as the one given in (9i), and the RR that inserts a y (cf. 10ii), is a rule which has independent motivation in the grammar, (cf. P&C and below):

(10) i) $[\underline{\quad}C]_N \rightarrow [\ [\underline{\quad}C]_N + ó\ ]_N$
           $\quad\quad$ <+m> $\quad\quad\quad$ <+m> $\quad$ <-m>

      ii) $\emptyset \rightarrow y\ /\ C\underline{\quad} + ó\ \#$

      where C is a fricative or an unexploded stop

The difference between (7i) and (7ii) is that the words in (7i) end in a stop, a nasal or a liquid, while the words in 7ii) end in a fricative.[3] It is, thus, the different phonological structures of the input words that determine the application of a specific RR. The two derivations are as follows:

(11)	LEXICON	wadad	kurus
	PFR	wadad+ó	kurus+ó
	RR	wadad+d+ó	kurus+y+ó[4]

The words in (7i) and (7ii) – contrary to the words in (1) – show what has been called "polarity," that is, while the input is masculine, the output is feminine. The syllabic structure of the input words has, thus, consequences also for the gender of the output.

Let us consider now the words listed in (7iii). Also in this case the PFR adds a stressed *o*, but the RR needed in order to arrive at the surface form is rather complex:

$$(12) \quad \check{V} \rightarrow \emptyset / \sigma1 \ [\begin{smallmatrix} r \\ l \end{smallmatrix} \ \underline{\quad} \ ]_{\sigma2} \ \sigma3$$
$$[-str]$$

where σ1 and σ2 are open syllables

(12) reads as follows: a short unstressed vowel is deleted when it occurs in the second syllable of a word of three syllables where the first and the second syllables are open and where the onset of the second syllable is a liquid.

Summing up so far, we have seen the following three types of plural formation:

(13)

	LEXICON	PFR	RR's	OUTPUT
i.	búug	→ á	RR 3	buugág
ii.	sánan	→ ó	RR 9	sanannó
iii.	kúrus	→ ó	RR 10	kurusyó
iv.	hílib	→ ó	RR 12	hilbó

As it can be seen by now, adopting the solution in (3ii) allows us to capture a generalization at the "morphological" level: the PFR always adds a (stressed) vowel.

Besides nouns ending in consonants, there are also masculine nouns ending in a vowel in Somali. Consider the following data:

(14)

i.	díbi	→	dibiyó	'bull'
	ri	→	riyó	'goat'
	falloóci	→	falloociyó	'irritable (person)'
ii.	gábay	→	gabayó	'poem'
	kúray	→	kurayó	'boy'
	fúlay	→	fulayó	'coward'
iii.	aabbé	→	abbayaál	'father'
	waarabé	→	waarabayaál	'hyena'
	baré	→	barayaál	'teacher'

Here the syllabic structure of the input word plays no role: the monosyllabic *ri* and the polysyllabic *falloci* form the plural in the same way. In (14i) there is again RR 10 at work (*y* insertion),[5] while in (14ii) no rule is needed. The situation is more complex in (14iii). A possible derivation is the following:

(15)	LEXICON	bare[6]
	PFR	bare+aal
	RR 10	bare+y+aal
	e → a	bara+y+aal (Vowel Harmony)

From what we have seen so far, plural formation of masculine nouns in Somali has a "center" of high regularity: addition of a stressed vowel (á for monosyllable, ó for polysyllables; only when the ending vowel is a stressed *e* the PFR attaches -*aal*), and a set of different RR's. The final shapes of words is determined by the RR's that apply after PFR. These RR's apply taking into account the gender, the phonological structure, and the syllabic structure of the input.

Let's now consider the plural formation of feminine nouns:

(16)

layr	→	layró	'air'
hal	→	haló	'camel (fem.)'
kaliíl	→	kaliiló	'hot season'
diín	→	diimó	'religion'
laáb	→	laabó	'brest'
dheg	→	dhegó	'ear'
seéf	→	seefó	'sword'
bah	→	bahó	'son' (of the same mother)

As it is evident, plural formation of feminine nouns is (morphologically) similar to plural formation of masculine nouns. The general rule is 'add *o*'. The difference lays, once again, in the various RR's needed. Let us consider some points.

a) The distinction between monosyllabic words and polisyllabic words is not relevant for feminine nouns (both *hal* and *kalifl* form the plural in the same way).

b) the final consonant of the input word, which plays an important role in the plural formation of masculine nouns is not relevant for feminine nouns (both *laab* – ending in a stop – and *seef* – ending in a fricative – form the plural in the same way).

Basically, we can say that PFRs require more subtle information to operate for masculine nouns than to operate for feminine nouns. This is confirmed also with regard to RR 12 seen above which is also needed for feminine nouns, but in a simplified form. Consider the following (feminine) nouns and their plural:

(17)   niríg   →   nirgó   'little camel (fem.)'
       ciríb   →   cirbó   'heel'
       habár   →   habró   'old person (fem.)'
       luqún   →   luqmó   'neck'
       gacán   →   gacmó   'hand'

For the first two examples the RR seems to be exactly the same as the one seen in (12). But in the last three examples we see that the deletion of the short vowel is not restricted by the condition that the onset of the second syllable should be a liquid. The rule operates with less restrictions.[7] The RR in question will be thus:

(18) $\check{V} \rightarrow \emptyset / \sigma 1 \ [\underline{\hspace{1cm}}]_{\sigma 2} \ \sigma 3$
      [-str]

Also feminine nouns can end in a vowel, as can be seen below:

(19)   i.   mindí   →   mindiyó   'knive'
          tacsí   →   tacsiyó   'condolence'

   ii.   runléy   →   runleeyó   'the one (fem.) who says the truth'
        beenléy   →   beenleeyó   'the one (fem.) who lies'

  iii.   asmó   →   asmoóyin   'curse'
        arrimisó   →   arrimisoóyin   'adviser (fem.)'

(19i) parallels (14i). (19ii) parallels (14ii) and is perfectly predictable. In (19iii) we find a different ending (*-in*) but the usual pattern, plus the lenghthening of the stressed vowel of the singular form.[8]

We can summarize what has been said so far in the following table:

(20)

Masc	Monosyll	Ending	Suffix	RR	Polarity
+	+	C	á	(á)Cα#	−
+	− 1.	C	ó	Cα(ó)#	+
+	− 2.	C	ó	y(ó)#	+
+	− 3.	C	ó	0C6#	−
+	1.	V -i	ó	y(ó)#	+
+	2.	V -ay	ó	----	+
+	3.	V -é	aál	y(ó)#	+
−	1.	C	ó	----	+
−	2.	C	ó	0C6#	+
−	1.	V -i	ó	y(ó)#	+
−	2.	V -ey	ó	----	+
−	3.	V -ó	in	y(in)#[9]	+

To summarize, we have seen that plural formation in Somali can be accounted for by the interaction of an inflection rule and a set of readjustment rules. The PFR is a very simple rule which, basically, attaches a stessed vowel to the input word. The RR's are basically of three kinds: copying, insertion and deletion rules. The RR's that apply to feminine nouns are similar to those that apply to masculine nouns, but they are simpler, in the sense that they need less information to operate.

NOTES

1. Cf. Scalise (1984).
2. The data discussed here are from Puglielli & Ciise (1984) (P&C hereafter). In this paper we will only consider "standard" plural formations and not other types of plural such as a) "pluralia tantum" (*hablo* 'girls'), b) plurals formed by suprasegmental means (cf. *túug* vs *tuúg* 'thief - tieves'), c) arabic plurals (*musbáar* vs. *musaabiúr* 'nail - nails'), d) 'plural of the plural' formations (*buugag-y-ow* 'groups of books) or e) reduplication (*laab* vs. *laaláab* 'curve - curves'). For more information about these formations, see P&C.
3. There are nouns ending in stops which form the plural with 'y+o'. For these cases, P&C suggest that the stop is 'unreleased'.
4. This derivation, as the others that will follow does not include for simplicity stress and the CAR mentioned before.

5. It should be noted that the context of this insertion rule is not the same as in (10). In (10) the rule inserts y between a consonant and *ó*, and in (14i) between a vowel and *ó*. This difference requires a reformulation of rule (10) such as

$$\emptyset \rightarrow y \; / \; \underline{\hspace{1cm}} \; ó \; \#$$

6. Actually, *baré* is a derived word from *bar+e*, where *e* is an agentive derivational suffix (cf. Puglielli 1984).

7. We will not discuss here the change $n \rightarrow m$, which is very common in the whole phonology of readjustment in Somali (cfr. *nin* → *nimán* 'man - men', *waran* → *warmo* 'lance - lances'.

8. Notice that *in* is not stressed, therefore lengthening is in a way the opposite of CAR; the addition of a syllable without stress results in the legthening of the stressed vowel.

9. Plus lengthening.

REFERENCES

Puglielli, A. (1984) "La derivazione nominale in somalo", *Studi somali*, 5, 3-52.
Puglielli, A. and Ciise M.S. (1984) "La flessione del nome", *Studi Somali*, 5, 53-112.
Scalise, S. (1984) *Generative Morphology*, Foris, Dordrecht.