On phonological processes in the "3rd conjugation" of Somali

Xavier Barillot & Philippe Ségéral
Université de Nice / Université Paris VII

The purpose of this paper is to show that, despite certain apparent discrepancies, the flexion of the so-called "third conjugation" (henceforth c3) in Somali is regular.[1] This regularity, we argue, can be captured assuming mainly that some consonants which are phonetically simple are geminated phonologically.

1. Introduction

Somali c3-verbs can be roughly described as middle voice verbs (Hayward 1975: 209-214, Saeed 1998: 124-127). Morphologically speaking, c3-verbs are derived by attaching one of two slightly different lexical suffixes (henceforth LS's) directly to the basic stem. In the imperative 2S, which is the most usual quotation form of verbs in Somali, the LS's are *-o* and *-so*[2], hence two subclasses, the "*-o* verbs" and the "*-so* -verbs" (3B and 3A, respectively, in Zorc, Osman & Luling 1991: xix and Orwin 1995: 82, 3A and 3B in Saeed 1998: 74). In fact, the LS's appear in this shape in the quotation form (imper. 2S) only. In all other forms various allomorphs can be observed. The exhaustive[3] set of these allomorphs is given in [1] below:

[1] a. "*-o* verbs" *-o, -t-, -d-, -at-, -ad-, -an-*
b. "*-so* verbs" *-so, -st-, -sat-, -sad-, -san-*

The part of this complex allomorphy in which an oral coronal consonant is involved (namely *-t-, -d-, -at-, -ad-* and *-st-, -sad-, -sat-*)[4] is the main focus of this paper and will be thoroughly discussed later. For the time being, let us assume that the LS's in question have the underlying forms */-at-/* and */-sat-/*[5].

The derived base, [Stem + LS], can take several other morphemes: 1. the Subject Agreement [henceforth SA] morphemes, 2. the Tense morphemes, 3. in the 2P and 3P persons, a Plural morpheme (*-n*). These morphemes are not specific to c3 ; they are common to the three conjugations. Let us consider them in the Habitual Present [2a] and the Past [2b] of the verb *keen* 'bring', a verb of c1:

[2]

	a. Habitual Present				b. Past			
1S	*keen*		*aa*		*keen*		*ay*	
2S	*keen*	*t*	*aa*		*keen*	*t*	*ay*	
3mS	*keen*		*aa*		*keen*		*ay*	
3fS	*keen*	*t*	*aa*		*keen*	*t*	*ay*	
1P	*keen*	*n*	*aa*		*keen*	*n*	*ay*	
2P	*keen*	*t*	*aa*	*n*	*keen*	*t*	*ee*	*n*
3P	*keen*		*aa*	*n*	*keen*		*ee*	*n*

In [2][6], for sake of clarity, the different morphemes have been separated and aligned. The Tense marker is *-aa-* in the Habitual Present and *-ay-/-ee-* in the Past. The SA markers are common to both tenses: *-t-* for 2S, 3fS and 2P, *-n-* for 1P, and

absence of an overt SA marking for 1S, 3mS and 3P. Because of its particular relevance in the following analysis, we draw the attention of the reader to the fact that the SA markers are either a single consonant (*t* or *n*), or null.

2. *a / ø* alternations

We can now turn to examine some examples of c3-verbs. Six of them are given (in the full paradigm of the Habitual Present) in [3]:

[3]	a. *qab-o*	b. *tuur-o*	c. *daqn-o*	d. *kabb-o*	e. *farii-so*	f. *yoob-so*
	'hold'	'pile up'	'feel pain'	'sip'	'sit down'	'gather'
1S	*qabtaa*	*tuurtaa*	*daqnadaa*	*kabbadaa*	*fariistaa*	*yoobsadaa*
2S	*qabataa*	*tuurataa*	*daqnataa*	*kabbataa*	*fariisataa*	*yoobsataa*
3mS	*qabtaa*	*tuurtaa*	*daqnadaa*	*kabbadaa*	*fariistaa*	*yoobsadaa*
3fS	*qabataa*	*tuurataa*	*daqnataa*	*kabbataa*	*fariisataa*	*yoobsataa*
1P	*qabannaa*	*tuurannaa*	*daqnannaa*	*kabbannaa*	*fariisannaa*	*yoobsannaa*
2P	*qabataan*	*tuurataan*	*daqnataan*	*kabbataan*	*fariisataan*	*yoobsataan*
3P	*qabtaan*	*tuurtaan*	*daqnadaan*	*kabbadaan*	*fariistaan*	*yoobsadaan*

As a first step, we will focus attention on 1S forms only.

Let us begin by considering one contrast existing between *daqnadaa* [3c] and *qabtaa* [3a]: while in the former, the LS displays a vowel ***a***, no vowel appears in the LS of the latter: *daqn-**a**d-aa vs qab-**ø**t-aa*. Further, *tuuro* [3b] and *fariiso* [3e] behave in this respect as *qabo*, while *kabbo* [3d] and *yoobso* [3f] behave like *daqno*.

Observe then that the presence of ***a*** in the LS correlates with the presence of a CC cluster to the immediate left of the alternation-site: *daqn- / kabb- / yoob-s-**a**d-aa*. Conversely, when one consonant only stands to the left of the LS, no vowel appears: *qab- / tuur- / fariis-**ø**t-aa*.

This *V/ø* alternation is not an isolated case in Somali. In order to correctly capture the regularity at work here, it will prove fruitful to briefly review another case of *V/ø* alternation, that can be observed for a subset of c1-verbs:

[4]	a. *gudub*	b. *kidif*	c. *boqor*	d. *dhereg*	e. *qahar*
	'cross'	'shop'	'make king'	'feel full'	'make trouble'
1S	*gudbaa*	*kidfaa*	*boqraa*	*dhergaa*	*qahraa*
2S	*gud**u**btaa*	*kid**i**ftaa*	*boq**o**rtaa*	*dher**e**gtaa*	*qah**a**rtaa*
3mS	*gudbaa*	*kidfaa*	*boqraa*	*dhergaa*	*qahraa*
3fS	*gud**u**btaa*	*kid**i**ftaa*	*boq**o**rtaa*	*dher**e**gtaa*	*qah**a**rtaa*
1P	*gud**u**bnaa*	*kid**i**fnaa*	*boq**o**rnaa*[7]	*dher**e**gnaa*	*qah**a**rnaa*[6]
2P	*gud**u**btaan*	*kid**i**ftaan*	*boq**o**rtaan*	*dher**e**gtaan*	*qah**a**rtaan*
3P	*gudbaan*	*kidfaan*	*boqraan*	*dhergaan*	*qahraan*

First of all, observe that in [4] the site of the alternation is no longer a LS but the stem itself. Then, two things are striking:

1. the second stem-vowel (in bold face in [4]) which appears in some forms is always identical to the first stem-vowel for each verb. This circumstance precludes any analysis involving a syncope process. Indeed, if the second stem-vowel were underlyingly present and was being deleted under certain conditions (Saeed 1998: 23), one could hardly account for the systematic identity of the two stem-vowels in those verbs. Rather, such verbs with stem-vowel alternation (henceforth SVA verbs)

are to be viewed as lacking a second stem-vowel, that is, they are underlyingly /CvCC-/: the second vowel that sometimes appears in the stem is nothing but a copy of the stem-vowel. The process of copy is triggered in certain syllabic situations (see below). Within this analysis, the systematic identity of the two vowels follows[8].

2. the copy of the first stem-vowel appears within the stem whenever the SA marker is a consonant, *-t-* (2S, 3fS, 2P) or *-n-* (1P), but is absent whenever the SA marker is null (1S, 3mS, 3P). Clearly, the copy of the stem-vowel is triggered in order to avoid a word-internal *CCC cluster: */gudb-ø-aa/ > gudbaa* (1S), but */gudb-t-aa/ > gud**u**btaa* (**gudbtaa*) (2S, 3fS).

Returning now to c3-verbs in [3], we can interpret the surfacing of the LS vowel as resulting from the same ban on internal *CCC clusters: we get *daqn**a**daa, kabb**a**daa, yoobs**a**daa* because of the impossibility of **daqntaa, *kabbtaa, *yoobstaa,* while *qabtaa, tuurtaa, fariistaa* are well formed.

Nevertheless, the LS *V/ø* alternation displayed by c3-verbs differs from the stem-internal alternation we observed in SVA c1-verbs in two respects:

1. the illicit cluster *CCC is broken into CC**v**C, not into C**v**CC: *yoobs**a**daa*, not **yoob**a**staa* (*vs gud**u**btaa*, not **gudb**u**taa*). In other words, the alternation site is immediately to the left of the C (*t/d/n*) of the LS

2. the inserted vowel cannot any longer be a copy of the stem-vowel, since it is always *a*, independently of the stem-vowel.

These two points suggest that the phonological shape of c3-LS's itself is responsible for the place and the nature of the appearing vowel. We propose that the underlying structure of the lexical c3-LS's involves a vocalic element *A*[9], and that this element is not associated to the prosodic structure unless the syllabic configuration requires it, that is, is lexically a floating segment. Indeed, if *A* was lexically linked to the prosodic structure, we should have to suppose a syncope process when it does not surface. That is, we should have to suppose two opposite processes at work in Somali *V/ø*-alternations : vowel propagation in SVA verbs (recall that the *V/ø* alternation *cannot* be due to a syncope process in this case) and vowel syncope in LS's. This is unlikely and undesirable. The LS's that characterise the c3-verbs are therefore to be represented as in [5a] and [5b]:

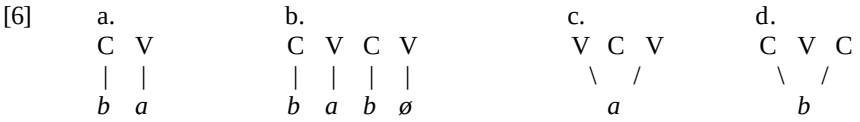
[5] a.

b.

where the consonantal elements *t* (and *s*) are linked to the prosodic structure of the LS's, while the vocalic one *A* is not, *i. e.* is floating.

The above representations in [5] are given within the syllabic framework we assume in this paper, *i. e.* the "CVCV" model as defined in Lowenstamm (1996). In this model, which refers more generally to Government Phonology, the syllabic inventory is maximally restricted to simple (= non-branching) Nuclei and simple Onsets, which monotonously alternate in the chain. We give in [6] the representations, in this frame, of familiar phonological objects such as a light syllable [6a], a closed syllable [6b], a long vowel [6c], a geminated consonant [6d]:

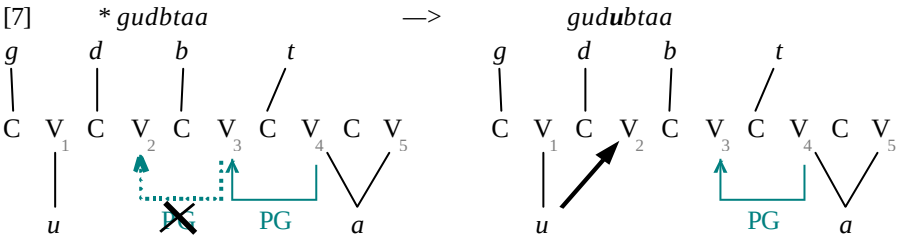
[6]



We stated above that the underlying *A* of the LS is associated in *yoobs**a**daa* and the stem-vowel is copied in *gud**u**btaa*, in order to avoid an illicit internal *CCC cluster. This association does not occur in *qabtaa* nor in *gudbaa*: here we only have a (licit) internal CC cluster. Observe that, according to [6], the forms just quoted are underlyingly */qab-øt-aa/* and */gudøb-aa/*, that is they each involve an empty nucleus, so that the following question arises: why in this case this nucleus may remain empty while the same nucleus must be filled in *yoobs-**a**d-aa* and *gud**u**b-t-aa*?

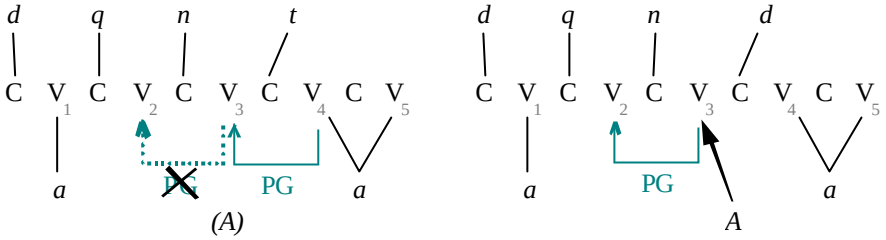
We claim that in all cases the presence or absence of a vowel on the alternation site can be captured as a mere application of the Empty Category Principle (henceforth ECP), a central device of Government Phonology (Kaye, Lowenstamm & Vergnaud 1990). The ECP defines under which condition an empty nucleus may remain such: an empty nucleus may remain empty only if it is properly governed. Proper Government (henceforth PG) is a form of leftward government holding between two adjacent nuclei; in order to properly govern the nucleus preceding it, a nucleus must not itself be properly governed, *i. e.* must not be empty (Kaye, Lowenstamm & Vergnaud 1990, Charette 1990, Scheer 1998). This in particular precludes any internal sequence of two empty nuclei. If such a situation arises in the course of the derivation, one of the two empty nuclei must be filled. In all cases where a vowel appears (*daqn**a**daa* or *gud**u**btaa*), such a conflicting sequence of two empty nuclei is underlyingly present: */daqønøtaa/*, */gudøbøtaa/*, while only one empty nucleus is present in the forms where no vowel appears on the alternation site: */qabøtaa/*, */gudøbaa/*. In the latter situation, nothing happens, since the empty nucleus is properly governed by the following one. In the former the conflict must be fixed: in the case of SAV verbs the syllabic conflict is repaired by propagating the stem-vowel toward the neighbouring position[10] (see V_2 in the example in [7]):

[7]



while in the case of c3-verbs, the same syllabic conflict is prevented by association of the floating vocalic element *A*, since this element is available in the lexical structure of the LS:

[8] * *daqntaa* —> *daqn**a**daa*



Summarising, we may say that for c3-verbs, the surfacing of the floating vowel *A* of the LS always results from a lack of PG at the underlying level.

Practically - and the reader's attention is drawn to this point, which will be revealed as crucial in what follows, *the presence of a vowel on the V/ø alternation site always correlates with the presence of a CC cluster, either to the left or to the right, or both*. (Note that one could think, considering 1S forms like *daqn**a**daa*, or *yoobs**a**daa*, that the association of the LS *A* takes place only when a *left* CC cluster is present. But 1P forms like *qab**a**nna*, *tuur**a**nnaa* or *fariis**a**nnaa* in [3] show clearly that the same happens when the CC cluster is to the *right* of the site. And, finally, in 1P forms *daqn**a**nnaa*, *kabb**a**nnaa*, *yoobs**a**nnaa*, the presence of *A* is triggered by *both* a left and a right CC cluster. Even in those forms indeed where a sequence of *three* empty nuclei arises (*/yoobøsønønaa/*), the association of the LS *A* yields a situation with no remaining lack of PG: in */yoobøsAnønaa/* the first empty nucleus is properly governed by the following *A*, the second by the final *aa*).

3. Voicing of *t*

So far, we have accounted for the *a / ø* alternations that can be observed in c3-verbs 1S forms. However this is not the only process that takes place in those forms: while the LS consonant is *t* in *qabtaa* or *tuurtaa*, it surfaces as *d* in *daqnadaa*, *kabbadaa* or *yoobsadaa*.

This is a mere instantiation of a very general sound rule of Somali according which *t* voices (and further spirantises) in intervocalic position. Nominal morphology provides an abundance of examples of this sound rule. The feminine determiner in Somali is */t + V/* (where *V* is either *a*, *u*, or *ii* depending on further grammatical conditions). Since in Somali the determiner is suffixed and feminine nouns end either in a consonant or in a vowel (*i* or *o*), one can verify that *t* remains unchanged when the noun ends in a consonant[11], while it becomes voiced whenever the noun ends in a vowel: */laf-ta/ > lafta* 'the bone', */naag-ta/ > naagta* 'the woman', */shimbir-ta/ > shimbirta* 'the bird', etc., but: */mindi-ta/ > mindi-da* 'the knife', */hooyo-ta/ > hooyada*[12] 'the mother'.

Though the velar consonant */k/* is not yet of direct concern, let us state here, because of the further relevance of the point, that the behaviour of */k/* wholly parallels that of */t/* in Somali. Indeed, */k/* voices in intervocalic position too, as is evidenced by the behaviour of the masculine determiner which is similar to the feminine one except that the consonant is *k* instead of *t*: */dab-ka/ > dabka* ' the fire', */beer-ka/ > beerka* 'the liver', */macallin-ka/ > macallinka* 'the teacher', but */guri-ka/ > guriga* 'the house', */qoraa-ka/ > qoraaga* 'the writer'[13]. Further evidence comes from the concatenation of preverbal Prepositions: */u + ka / > uga* 'to + from', or concatenation of Object Pronouns with Prepositions : */i + ku/ > igu* 'me + in'.

Since voicing of *t* and *k* happens in Coda-position too, the two relevant sound rules can be stated as in [9][14]:

[9] a. */t/* > *[d]* / V ___
b. */k/* > *[g]* / V ___

The rule [9a] accounts for all cases where the allomorph *-ad* of the c3 LS appears instead of *-at*, among which 1S forms of [3].

4. Intervocalic *t*

We turn now to the 2S forms in [3] which are repeated in [10] for convenience (together with the 1S forms we have just dealt with):

[10]	a. *qab-o*	b. *tuur-o*	c. *daqn-o*	d. *kabb-o*	e. *farii-so*	f. *yoob-so*
	'hold'	'pile up'	'feel pain'	'sip'	'sit down'	'gather'
1S	*qabtaa*	*tuurtaa*	*daqnadaa*	*kabbadaa*	*fariistaa*	*yoobsadaa*
2S	*qabataa*	*tuurataa*	*daqnataa*	*kabbataa*	*fariisataa*	*yoobsataa*

All of these 2S forms overtly infringe [9a], since all display an intervocalic unvoiced *t*. Furthermore, this is not the only anomaly that these forms display. In *daqn**a**taa*, *kabb**a**taa* or *yoobs**a**taa*, the presence of the LS vowel ***a*** is now for us of no surprise : in all cases, there stands a CC cluster to the left of the alternation site and therefore the surfacing of the LS floating vowel is just what is expected. But this is not the case for *qab**a**taa*, *tuur**a**taa*, *fariis**a**taa*. In these verbs, indeed, only one *[t]* stands to the left of the alternation site.

Faced with such a situation, we can either renounce the generalisations made above, or, still maintaining them, we proceed to the assumption that the *t* standing to the right of the alternation site in these forms must be a CC cluster. This conclusion is confirmed by the morphological analysis. There are *two* different *t*'s in the underlying 2S forms of c3-verbs: the one of the LS (*-(A)t-*) and the one of the SA *-t-* of 2S: *qabataa* < */qab-(A)t-t-aa*.

This leads to the claim that in Somali a simple intervocalic *[t]* is the regular interpretation at the phonetic level of a geminated underlying */tt/*. Under this assumption,

1. the triggering of ***a*** on the *V/ø* alternation site is regular. According to [6d], the two C positions identified by the geminate consonant straddle an empty V position. Being empty, this nucleus is unable to properly govern the nucleus preceding it: this lack of PG triggers the association of the floating *A* of the LS exactly as shown in [8],

2. moreover, the resistance to voicing just pointed out above becomes natural: typically geminate consonants resist weakening processes such as voicing, spirantization, or even disappearance, that single consonants are likely to undergo in intervocalic position (Perlmutter 1995: 309).[15]

Note that, if a single phonetic consonant may represent an underlying geminate, one could instead propose that it is the *b* in *qab**a**taa* or *r* in *tuur**a**taa* which is the CC cluster which forces the presence of the LS ***a***. But 1. the morphological analysis does not fit with this hypothesis and 2. if *b* or *r* represented a CC cluster (*i. e.* were geminated), this would be phonetically audible: indeed, both *b* and *r* belong to the set of Somali consonants which may phonetically geminate.[16]

In fact, it is precisely the existence in the language of a general rule of voicing of *t* in intervocalic position and the *V/ø* alternation regularity which allow phonologically geminate */tt/* not to be interpreted as such phonetically: the recoverability of the phonological geminate is in any case guaranteed by these external clues. In the same way, phonetic intervocalic voiced geminates like *[bb]* are optionally[17] pronounced as a simple stop *[b]* as already noted in Saeed (1982: 4). This possibility follows from the existence in the language of a rule of spirantisation of intervocalic simple voiced stops: an intervocalic *[b]* cannot be therefore the interpretation of a single intervocalic */b/* but only of a spirantisation-resistant phonological geminate. In intervocalic position, the lack of spirantisation is thus sufficient to manifest the underlying geminate.

Further, several puzzling facts in Somali cannot be accounted for unless it is assumed that an intervocalic *[t]* is the phonetic interpretation of an underlying geminate. For instance, it has been noted (Orwin 1995: 75, Saeed 1998: 23) that among potentially SVA c1-verbs, none of the verbs with *t* as medial consonant display the expected *V/ø* alternation. Contrasting with *gud**u**b / gudbaa* (see [2]), indeed, we have:

[11]	Imper. 2S	Pres. 1S	
	*mat**a**g*	*mat**a**gaa (*madgaa)*	'vomit'
	*hit**i**q*	*hit**i**qaa*	'walk slowly'
	*qot**o**n*	*qot**o**maa*	'be upright'[18]
	*mut**u**x*	*mut**u**xaa*	'speak without reserve'

The presence of the second (copied) vowel in all forms of these verbs cannot be accounted for by the impossibility for */t/* of voicing to *[d]* if this is then syllable-final, as proposed in Saeed (1998: 23). The verb *gunud* 'tie a knot' has a final underlying */t/* as is evidenced by the Pres. 1S *guntaa*. Now this */t/* results in *[d]* in *gunud* and yet it is syllable-final, and the same, word-internally, in */gunut-n-aa/* > *gunudnaa*. If the medial *[t]* of *matag / matagaa* is a geminate, an alternative explanation is possible. Indeed the stem-vowel copy will always be triggered in such a verb, since the empty nucleus the two members of the geminate straddle needs always to be properly governed by the nucleus following it. Moreover the lack of voicing of *[t]*, which otherwise remains unexplained, follows.

In the same way, Somali, as pointed out in Keenadiid (1976: XXII), has a free variation between a geminate and a [nasal + obstruent] cluster as shown by the doublets *cadhdho* (usual spelling *cadho*) / *candho* 'scabies', *middi / mindi* 'knife', *higgo / hingo* 'hiccough', etc. Parallel cases involving *t*, such as *mitid / mintid* 'persevere', *cutub / cuntub* 'small group', *sitaacso / sintaacso* 'put ornamental trappings on a horse', *mantag / matag* 'vomit', etc. clearly show that an intervocalic *t* stands for a geminate.

A similar argument can be made in compounds. When the first member of a compound ends in a vowel, the first consonant of the second member is regularly geminated: *dhego* 'ears' + *beel* 'lack of' > *dhegabbeel* 'deafness'. But the same word *dhego* 'ears' + *tir* 'cancel' > *dhegatir* 'ignore, refuse to listen'. *t*, phonetically single but resisting voicing, behave in all as a geminate.

Now, let us turn to the rather spectacular case of c3-verbs ending in *-Vdo* such as *bado* 'increase (intr.)', *feydo* 'strip, undress', *gaado* 'choose', *gado* 'buy', *liido*

'be stupid / weak', *wado* 'drive', *qodo* 'dig', *quudo* 'eat, feed on', etc. See the 1S and 2S forms of one of these verbs in [12]:

[12]	imper. 2S	Pres. 1S	Pres. 2S	
	qaado	*qaataa*	*qaad**a**taa*	'take for oneself'

The Pres. 2S form, *qaad**a**taa,* is in no way different from the one of all c3-verbs (*cf.* [10]). In the 1S, exactly as in in *qabtaa* in [10], the LS vowel remains floating since no lack of PG arises, but the stem-final coronal[19] plus the LS *t* result in a single intervocalic unvoiced *[t]*. This is a remarkable argument in favour of the phonologically geminated status of an intervocalic *[t]*.

Finally, let us mention that such "virtual" geminates are not specific to Somali (see Lowenstamm 1996, and Scheer & Ségéral to appear b). For instance, it has long been argued (Leslau 1948) that in Chaha, a Semitic language spoken in Southern Ethiopia (Gurage subgroup), an important subset of root-medial single voiceless obstruents are the regular reflexes of geminated voiced ones in cognate languages. Furthermore, in Chaha itself, root-medial voiceless obstruents of that particular group regularly alternate with their voiced or spirantised counterparts (*e. g.* perf. 3mS *bätäräm,* juss. 3mS *yä-ßdär* 'be first, advance', perf. 3mS *mäkäräm,* juss. 3mS *yä-mgär* 'suppurate'). The most enlightening analyses of these highly complex alternations (see Petros 1997, 2000) involve a level of representation where the relevant obstruents are represented as underlying geminates, even though they surface as single phonetic segments.

5. Intervocalic *k, sh, w*

Let us consider now the case of c3-verbs ending in:

1. *-Vko*: *buko* 'become sick, *dhako* 'hide oneself', *tuko* 'pray', etc.
2. *-Vsho*: *cesho* 'keep', *faasho* 'predict', *gasho* 'wear', *tasho* 'consider', etc.
3. *-Vwo*: *daawo* 'watch', *dhaawo* 'hold a grudge', *duwo* 'bypass', etc.

all of which follow the same paradigm. One example of each type is given in [13], in the 1S form only, since this is sufficient for the sake of discussion:

[13]		a. *tuk-o*	b. *cesh-o*	c. *daaw-o*
		'pray'	'keep'	'watch'
	1S	*tuk**a**daa*	*cesh**a**daa*	*daaw**a**daa*

In all these 1S forms, the right-hand context of the alternation site undoubtedly is a single */t/*. This follows, first of all, from the morphological analysis: only one *t* is involved here, the one of the LS, since the SA marker is null in 1S forms. Moreover, this underlying */t/* does voice in intervocalic position, according to [9a].

The left context is as well a single consonant in all verbs, respectively *[k], [sh], [w]* (recall 'sh' is only a digraph for the voiceless palato-alveolar fricative). But, under the same reasoning as previously, the presence of the LS ***a*** in *tuk**a**daa, cesh**a**daa, daaw**a**daa,* forces us to claim that, despite their phonetically single character, *k, sh* and *w* behave as a CC cluster, that is are phonological geminates. If they were not, the LS vowel should not surface, since no CC cluster should stand neither to the right nor to the left of the alternation site, and the forms should be **tugtaa, *ceshtaa, *daawdaa,* paralleling those of *qabo* or *tuuro* in [10a-b], *qabtaa* and *tuurtaa*. It follows that, as any intervocalic *t*, any intervocalic *k, sh, w* should be, at the phonological level, a geminate in Somali.

As for *[k]*, recall that a single */k/* voices in intervocalic position as stated in [9b]. Yet, the intervocalic *k* in *tuko* or *tukadaa* resists voicing. That is, in a form like *tuk**a**daa* the situation wholly parallels the one we saw for *t*: the unexpected LS vowel surfacing *and* the unexpected resistance to voicing cannot be captured unless *[k]* is construed as a geminate at the phonological level.

We therefore claim that all intervocalic *[k], [sh], [w]* represent a geminate in Somali, exactly as *[t]* does. This is not only shown by the c3 facts discussed above, but moreover can be evidenced through the correctness of various predictions it entails. For instance, it predicts that any potentially SVA c1-verb with a medial *k, sh,* or *w* should behave exactly as similar verbs with medial *t* such as *matag* we saw previously (*cf.* [11]), *i. e.* that no *V/ø* alternation should take place in such verbs, *i. e.* the stem-vowel copy should always be triggered. This is always the case, see examples in [14]:

[14]		Imper. 2S	Pres. 1S		
	a. *k*	*feker*	*fek**e**raa*	**fekraa*	'think'
		makal	*mak**a**laa*		'put in danger'
		xukun	*xuk**u**maa*		'govern, rule'
	b. *sh*	*bishil*	*bish**i**laa*		'wedge'
		fashal	*fash**a**laa*		'fail'
	c. *w*	*cawar*	*caw**a**raa*		'hurt, wound'
		sawax	*saw**a**xaa*		'shout'
		tawal	*taw**a**laa*		'worry'

Another prediction is that, comparable to the free variation *nt / t* we mentioned in section 4, one should find as well *nk / k, nsh / sh, nw / w* free variations. The prediction is correct as is shown by the following doublets: *sonkor / sokor* 'sugar', *maankaal / maakaal* 'residue left after the pressing of sesame', *canshuur / cashuur* 'tax', *xaashi / xaanshi* 'paper, document', *filanwaa / filawaa* 'unexpectedly'.

Further, various independent pieces of evidence of the geminate character of *[k], [sh], [w]* can be put forth in Somali morphophonology. As for *[k]*, compare for instance the forms of the two c1-verbs *bug* and *cab* in [15]:

[15]			1S / 3mS	2S / 3fS	1P	2P	3P
a.	*bug*	'be sick'	*bukaa*	*bugtaa*	*bugnaa*	*bugtaan*	*bukaan*
b.	*cab*	'drink'	*cabbaa*	*cabtaa*	*cabnaa*	*cabtaan*	*cabbaan*

The stem-alternation these two verbs display is in itself puzzling. It is not our purpose here to address this point since it falls far beyond the scope of this paper. We just aim at pointing out that *k* relates to *g* in the forms of *bug* exactly as *bb* to *b* in those of *cab*. Now the contrast in this latter verb is undoubtedly geminate *vs* simple. The intervocalic *k* in *bukaa* therefore, though phonetically single, stands for a geminate. Note that it resists voicing.

In the case of *[sh]*, a striking evidence comes from the peculiar paradigm of SVA 1-verbs ending in *-l*, such as *qosol* 'laugh', *maqal* 'hear', *faxal* 'plant', *hadal* 'speak', etc. The paradigm of *maqal*, for instance, is given in [16] (in the Present):

[16]	1S / 3mS	2S / 3fS	1P	2P	3P
	maqlaa	*maq**a**sha*	*maq**a**lnaa*	*maq**a**shaan*	*maqlaan*

The stem-vowel is copied in 2S, 3fS, 1P and 2P forms, that is whenever the SA marker is not null, *i. e.* exactly as in all SVA verbs in [4]. However, superficially this should not be the case neither in 2S / 3fS *maq**a**shaa* nor in 2P *maq**a**shaan*: to the right of the alternation site there stands only one single consonant, *sh*. And yet the alternating ***a*** is present. Once again, we are forced to view a single consonant (*sh*) as a cluster, since it triggers the stem-vowel copy, exactly as the cluster *ln* in 1P *maq**a**lnaa*. Now, the attentive reader should have noticed that this *sh* is in itself intriguing, since neither is there any *sh* in *maqal* nor any *sh* in the set of SA markers. In fact, *sh* stands here for *two* consonants: the stem-final *-l* of *maqal* and the SA marker *-t-*. In Somali, indeed, in verbal as well as nominal inflection, any underlying sequence *l-t* results in *sh*: */bil-ta/ > bisha* 'the month', */duul-t-aa-n/ > duushaan* 'you [P] attack', etc. The geminate nature of intervocalic *sh* we claimed after analysing c3 -*Vso* verbs is therefore supported by the phenomena displayed by SVA c1-verbs ending in *-l*.

Finally, let us add that, conversely, any c3-verb ending in *-VCo* where C is not {*k, sh,w* }, but belongs to the following set {*b, d, g, dh, q, l, m, n, r, f, s, kh, ', c, x, h, y*}[20], behaves as *qabo* [3a]. Under our general conception, this means that the consonants of this set are not likely to be virtually geminated in intervocalic position.

6. Conclusion

Somali c3 finally appears as a uniform and regular paradigm. The case of LS allomorphs in *-o* and in *-n* aside (see note 4), all apparent complexities displayed by c3 are due to the behaviour of the LS's. As we demonstrated above, the behaviour of the LS's (in other words the distribution of the *-t, -d-, -at-, -ad-* and *-st-, -sad-, -sat-* allomorphs) can be comprehensively captured given the following assumptions:

1. any *V/ø* alternation is controlled by the Empty Category Principle
2. single underlying stops */t/* and */k/* become voiced in intervocalic position
3. some underlying geminates (namely geminated *t, k, sh, w*) are interpreted on the surface as phonetically single consonants

Unless this latter point is admitted, not only it is impossible to account for the distribution of allomorphs in c3, but moreover a large set of independent aspects of Somali morphology cannot be accounted for.

The c3 paradigm in Somali stands as a main piece of evidence that phonological gemination, that is the association of a consonantal segment to two skeletal positions, is not necessarily interpreted on the surface as a phonetic geminate.

Notes

1 The "first" conjugation (henceforth c1) is the simple conjugation, the "second" is a (generally) causative conjugation which is characterised by the presence of a lexical affix *-i(y)-/-is-/-in* attached to the stem.

2 The suffix *-so* can be analysed as combining two different morphemes: *-s-* and *-o*, the latter being properly the "middle-voice" morpheme, the former probably what remains of the "causative" morpheme — see note 1— (Saeed 1998: 74, 124).

3 We neglect the variants {*-od, -ot, -on*} that appear in two verbs, *soco*, 'walk' and *noqo* 'return', probably due to an harmonic process originating in the stem-vowel *o* and favoured by the "transparency" of the guttural stem-final consonants *c* and *q*.

4 No attempt will be made in the following pages to account for either o-allomorphs (*-o, -so*), or nasal ones (*-an-, -san-*). Indeed, in both cases, the allomorphy cannot be seemingly accounted for on phonological grounds (Saeed 1998: 124). In the case of the allomorphs involving an *n*, one could imagine that an assimilation process *tn* > *nn* takes place in 1P forms like *qabannaa* (see paradigms in [3]), since the LS consonant is immediately followed by the *n* of the 1P Agr. marker. Unfortunately, 1. the hypothesis comes up against the absence of such an assimilation in *gunudnaa* (**gununnaa*) < */gunut-n-aa/* 'tie a knot (1P)', 2. the allomorphs *-an- / -san-* appear even if no subsequent *n* is involved: as in the "infinitive" *qaban* and in all progressive forms *qab-an-ay-aa* (1S / 3mS), *qab-an-ay-s-aa* (2S / 3fS), *qab-an-ay-n-aa* (1P), etc.

5 Such a morpheme, involving a coronal consonant and conveying a general medio-passive meaning, is widespread in Eastern-Cushitic : see Hayward (1975) for comparision and reconstruction of its original form.

6 We use Somali orthography. Signs used correspond to those of the IPA but: *dh* = voiced retroflex stop, *x* = voiceless pharyngeal fricative, *c* = voiced pharyngeal fricative, *sh* = voiceless palato-alveolar fricative, *kh* = voiceless velar fricative, *j* = palato-alveolar affricate, *y* = palatal glide, *'* = glottal stop. Finally, vocalic length is noted by the doubling of the vowel.

7 Progressive assimilations *rn* > *rr* and *ln* > *ll* optionally take place in 1P; *e. g.* in [4] *boqornaa / boqorraa*, *qaharnaa /qaharraa* and, in [16], *maqalnaa / maqallaa*.

8 For a detailed discussion of "vowel-syncope" in Somali, see Barillot (1997) and Barillot & Ségéral in prep. For similar cases of vocalic copy in order to avoid prohibited consonantal clusters, see Clements (1993: 130-133) and Clements & Hume (1995: 260-261) on Kolami, or Ségéral (2000: 279) on Akkadian.

9 For sake of clarity, we make use of a capital letter to designate the inherent floating vowel of c3-LS. On the other hand, this design refers as well to the Theory of Elements (Kaye, Lowenstamm & Vergnaud 1985).

10 The same process occurs when the sequence of two empty nuclei is word-final. In the imper. 2S (quotation form), */gudøbø/*, the final nucleus, being empty, cannot properly govern the preceding one. It is filled by the stem-vowel copy: *gud**u**b*, not **gudb*.

11 However, 1. assimilation processes take place whenever the preceding consonant is *d* or *dh* (*d-t* > *dd* and *dh-t* > *dhdh*) and 2. voicing applies too whenever the preceding consonant belongs to the set of "guttural" consonants {*', h, x, c, kh, q*}.

12 A further process regularly changes stem-final *-o* to *-a* (see Saeed 1998: 26-27).

[13] */k/ > [h] / {e, o}__: /bare-ka/ > baraha* 'the teacher'. This is a baffling discrepancy, but rather than challenging the rule of intervocalic voicing of *k*, it should be left for further investigation (see Cardona 1981: 19-20).

[14] Voicing of */t/* and */k/* regularly applies not only in intervocalic position but in Coda-position too, either word-internal or word-final, *i. e.* in any syllable-final position (see Saeed 1998: 23). Therefore, the right context of the rule, since it is either __V, or __C, or __#, that is to say any possible context, is irrelevant: voicing of */t/* and */k/* is more simply a postvocalic process. What is specific of intervocalic position is that it triggers a further spirantisation of voiced stops. As a result, note that */t/* and */k/* appear as *[t]*, *[k]* only in 1. word-initial position and 2. in postconsonantal position, that is the cross-linguistic "coda-mirror" context carried out in Scheer & Ségéral (to appear a). For the case of underlying geminates */tt/* and */kk/*, see below.

[15] It could be adduced (and it has been, more or less, see Puglielli 1984: 24, Saeed 1998: 29) that the surface resulting forms are correctly derived if one supposes three (crucially) ordered rules in the grammar: (1) a rule of vowel deletion applying in a sequence of light syllables by which the LS *a* would be deleted in */qabataa/*, but neither in */joogsataa/* nor in */joogsattaa/*, (2) a rule of voicing of *t* in intervocalic position which would apply to */joogsataa/* but neither to */qabtaa/* nor */joogsattaa/*, and (3) a rule of degemination of */tt/* which would apply to */joogsattaa/*: The outputs would be, correctly, *[qabtaa]*, *[joogsadaa]*, *[joogsataa]*. However, we think this is to be rejected for two reasons: 1. on a general theoretical ground, because a similar set of ordered rules could be built as well if the facts were different or even the reverse; in other words because no naturality is involved (nor even required) in this way of accounting for the facts, 2. on a technical point, because Somali totally lacks any (word-internal) degemination process.

[16] Only *b, d, dh, g, l, m, n, r* can phonetically geminate in Somali (Orwin 1995: 6).

[17] This pronunciation was regular in the case of our informant Nuur Bashiir Cismaan Keenadiid, born in the North of Somalia. In his words, the *[b]* he pronounces in *[aabe]* for *aabbe* 'father' is a "strong b", while the one he pronounces as a bilabial voiced spirant *[ß]* in *gaaban* 'short' is "soft".

[18] */m/ -> [n] / __{#, C}* (Saeed 1998: 22). Hence the imper. 2S *qoton.*

[19] Saeed (1993: 55 and 1998: 29-30) rightly points out that the 2S of the c1-verb *qaad* is *qaaddaa*, while the 1S of the (derived from the former) c3-verb *qaado* is *qaataa*. Yet the same, seemingly, underlying sequence *d-t* is involved in both forms: */qaad-t-aa/* where *t* is the 2S SA marker for *qaaddaa*, and */qaad-t-aa/* where *t* is that of the c3-LS for *qaataa*. That is, the geminate resulting from *d-t* goes rightward in the former case (and as in feminine determiner suffixation, see note 11), and leftward in the latter. Saeed (1993: 55) concludes that "the sound changes are sensitive to whether the *t* in the verb is an AGR affix or the remnant of the LEX autobenefactive affix". We cannot in the limits of this paper address this puzzling point. But whatever the account to be given to this discrepancy is, it remains that the resulting intervocalic *[t]* in *qaataa* as well as the *[dd]* in *qaaddaa* represent two underlying consonants, that is are both a geminate.

[20] To our knowledge, there are no c3-verbs ending in *-Vjo.* Nevertheless, on the basis of other observations, it can be shown that *j*, as {*t, k, w, sh*}, behaves regularly as a geminate in intervocalic position (see Barillot & Ségéral in prep., and Bendjaballah, this volume). There is no verb ending in *-Vto.*

References

Agostini, Francisco, Annarita Puglielli, Ciise Moxamed Siyaad & *al*. 1985. *Dizionario Somalo-Italiano.* Roma: Gangemi.

Barillot, Xavier 1997. *Alternances "voyelle-zéro" et gémination consonantique en somali*. DEA thesis, Université Paris 7.

Barillot, Xavier & Philippe Ségéral in prep. "Alternances voyelles-zéro en somali."

Bendjaballah, Sabrina (this volume). "On Palatalisation in Somali."

Cardona, Giorgio Raimondo 1981. "Profilo fonologico del Somalo." *Studi Somali. Fonologia e lessico* ed. by Giorgio Raimondo Cardona & Francisco Agostino, 1: 3-26.

Charette, Monik 1990. "Licence to Govern." *Phonology Yearbook* 7: 223-253.

Clements, G. N. 1993. "Lieu d'articulation des consonnes et des voyelles: une théorie unifiée." *L'architecture des représentations phonologiques* ed. by Bernard Laks & Annie Rialland, 101-145. Paris: CNRS éditions.

Clements, G. N. & Elizabeth V. Hume 1995. "Internal Organization of Speech Sounds." *The handbook of Phonological Theory* ed. by John A. Goldsmith, 245-306. Oxford: Blackwell.

Hayward, R. J. 1975. "Middle Voice Forms in Eastern Cushitic." *Transactions of the Philological Society* 75: 203-224.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud 1985. "The Internal structure of Phonological Elements. A Theory of Charm and Government." *Phonology Yearbook* 2: 305-328.

Kaye, Jonathan, Jean Lowenstamm & Jean-Roger Vergnaud 1990. "Constituent structure and Government in Phonology." *Phonology Yearbook* 7: 193-231.

Keenadiid, Yaasiin Cismaan 1976. *Qaamuuska Af-Soomaaliga.* [Xamar-Firenze].

Leslau, Wolf 1948. "Le problème de la gémination du verbe tchaha." *Word* 4: 42-47.

Lowenstamm, Jean 1996. "CV as the only syllable type." *Current Trends in Phonology* ed. by Jacques Durand & Bernard Laks, vol. 2: 419-441. Salford: ESRI.

Orwin, Martin 1995. *Colloquial Somali. A Complete Language Course.* New York / London: Routledge.

Perlmutter, David 1995. "Phonological Quantity and Multiple Association." *The handbook of Phonologhical Theory* ed. by John A. Goldsmith, 307-317. Oxford: Blackwell.

Petros, Degif 1997. *Le système phonologique du chaha*. Doctoral dissertation, Montréal (UQAM).

Petros, Degif 2000. *Sound Mutations. The morphophonology of Chaha*. Philadelphia / Amsterdam: John Benjamins.

Puglielli, Annarita 1984. "La derivazione nominale in somalo." *Studi Somali. Aspetti morfologici, lessicali e della focalizzazione*, ed. by Annarita Puglielli, 5: 1-52..

Saeed, John Ibrahim 1982. "Central Somali. A grammatical Outline." *Afroasiatic Linguistics* 8.2: 1-43.

Saeed, John Ibrahim 1993. *Somali Reference Grammar* [2nd Revised Ed.]. Kensington, Maryland: Dunwoody Press.

Saeed, John Ibrahim 1998. *Somali*. Dublin: John Benjamins.

Scheer, Tobias 1998. "A unified model of Proper Government." *The Linguistic Review* 15: 41-67.

Scheer, Tobias & Philippe Ségéral to appear (a) "La Coda-miroir". To appear in *Bulletin de la Société de Linguistique de Paris* 96 (2001).

Scheer, Tobias & Philippe Ségéral to appear (b). "Abstractness in phonology: the case of virtual geminates". To appear in *Constraints and Preferences* ed. by Katarzyna Dziubalska-Kolaczyk. Berlin: Mouton / de Gruyter.

Ségéral, Philippe 2000. "Théorie de l'apophonie et organisation des schèmes en sémitique." *Research in Afroasiatic Grammar* ed. by Jacqueline Lecarme, Jean Lowenstamm & Ur Schlonsky, 263-299. Amsterdam / Philadelphia: John Benjamins.

Zorc, R. David, Madina M. Osman & Virginia Luling 1991. *Somali - English Dictionary* [2nd Revised and Expanded Ed.]. Kensington, Maryland: Dunwoody Press.