J. Linguistics (2017), 1–40. © Cambridge University Press 2017
doi:10.1017/S002222671700024X

Verbal classes in Somali: Allomorphy has no classificatory function[1]

XAVIER BARILLOT

Université Côte d'Azur, CNRS, BCL

SABRINA BENDJABALLAH

Laboratoire de Linguistique de Nantes, UMR 6310 CNRS & Université de Nantes

NICOLA LAMPITELLI

Université François Rabelais de Tours & Laboratoire Ligérien de Linguistique UMR 7270 CNRS

(Received 8 December 2015; revised 24 March 2017)

This paper focuses on the complex derivational and inflectional morphology of Somali (East Cushitic) verbs. Somali verbs are traditionally cast in three major classes, depending on specific lexical suffixes (Saeed 1993). It is assumed that these classes must be distinguished because the relevant suffixes trigger a morphologically conditioned allomorphy. We argue against this view and claim that the allomorphic patterns targeting each class are epiphenomenal. Our analysis, couched within the theoretical framework of Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990) and the CV-model (Lowenstamm 1996), shows that the allomorphy in question is in fact phonologically conditioned. In particular, we establish unified representations of the two major lexical suffixes – the causative and the autobenefactive – and claim that all surface realizations of these markers result from the application of regular phonological rules. Thus, contrary to what appears at first sight, Somali displays a single verbal class whose three subclasses are phonologically (not morphologically) defined.

[1] We thank the anonymous *Journal of Linguistics* referees for their comments and remarks on previous versions of this paper. We also wish to thank Philippe Ségéral for having awoken our interest in Somali phonology in a fieldwork seminar at Paris 7, with Bashiir Nur Keenadiid as informant, as well as Elsa Godon and David Le Gac as participants.

We adopt the following abbreviations throughout the article: 1P, 2P, 3P = first, second, third person plural; 1S, 2S, 3S = first, second, third person singular; 3FS = third person feminine singular; 3MS = third person masculine singular; AUTOBEN = autobenefactive; CAUS = causative; DET = determiner; DITR = ditransitive; F = feminine gender; IMP = imperative; INTR = intransitive; LEX = lexical suffix; LIC = licensing; M = masculine gender; N = noun; NUM = number; PG = proper government; PL = plural; PNG = person, number, gender; PRES = present; SG = singular; TAM = tense, aspect, mood; TR = transitive; o.s. = 'oneself'; s.o. = 'someone'; s.t. = 'something'. For the transcriptions, we use IPA 2005 symbols.

KEYWORDS: allomorphy, Cushitic, Government Phonology, palatalization, Somali, verbal morphology

1. INTRODUCTION

Allomorphy, roughly defined as the situation in which a given feature is spelled out as two or more distinct exponents (see Aronoff 1976, Anderson 1992, Embick 2010, Bonet & Harbour 2012, Trommer 2012 among many others), remains an intriguing phenomenon in many theoretical frameworks. Although morphological theories propose radically different approaches to allomorphy, there is a general consensus on the fact that the systematic vs. the unsystematic application of a phonological rule plays a crucial role in determining the limits of this phenomenon (see Kiparsky 1996). Thus, the traditional view distinguishes between phonologically conditioned allomorphy and morphologically conditioned allomorphy. Phonologically conditioned allomorphy is predictable from the phonological rules of the language while morphologically conditioned allomorphy is not.

Somali, an Afroasiatic language belonging to the Lowland Eastern branch of the Cushitic group, is characterized by a rich inflectional morphology in both its nominal and verbal systems. Nouns are divided in various inflectional classes on the basis of a few morpho-phonological parameters (Andrzejewski 1964, 1979; Puglielli & Siyaad 1984; Banti 1988; Saeed 1993, 1999; Orwin 1995; Godon 1998). Verbs are traditionally cast in three main conjugations (Saeed 1993, Orwin 1995) plus a fourth conjugation, also known as the 'hybrid conjugation' (Andrzejewski 1969, Banti 2012, Puglielli & Mansuur 2012). These conjugations correspond to different verb classes, characterized by specific suffixes (e.g. causative, autobenefactive). It is assumed that these conjugations must be distinguished because the relevant suffixes trigger a morphologically conditioned allomorphy. In other words, allomorphy is taken to have a central role in the Somali verb morphology: it has a classificatory function, leading to the distinction of three main inflectional classes.

In this paper, we argue against this view. More precisely, we show that alleged cases of morphologically conditioned allomorphy can be reduced to phonologically conditioned allomorphy: surface forms result from the application of regular phonological rules to underlying structures. As a consequence, the notion of conjugation can be dispensed with. This analysis relies on a careful examination of the internal phonological structure of the relevant verbal suffixes. Endorsing an approach that decomposes the morphemes 'all the way down', we propose articulated phonological representations, and show how general phonological rules apply to these representations.

Our analysis shows that abstract phonological representations have a direct advantage on the way allomorphy can be analyzed. It also has a consequence on the status of 'paradigms' as defined by wide morphological literature (Aronoff 1994; Carstairs-McCarthy 1998, 2005; Blevins 2006, 2015; and references

therein). Current morphological theories discuss whether paradigms are active linguistic items. We participate in this debate in showing that the notion of paradigm is not relevant in the Somali verb system: contrarily to what appears at first sight, Somali displays a unique verbal conjugation.

Our work relies on data taken from the relevant standard literature as well as our own fieldwork with native speakers of Somali who have French as their second language. The analysis is couched within the general theoretical framework of Autosegmental phonology (Goldsmith 1978) and its particular versions known as Government Phonology (Kaye, Lowenstamm & Vergnaud 1985, 1990) and CV-phonology (Lowenstamm 1996).

The article is organized as follows: Section 2 introduces Somali verbal morphology as described in the literature. Sections 3 and 4 analyze conjugations 2a/b and 3a/b, respectively. We show that there is no need to consider each conjugation as a separate inflectional group. In Section 5, we conclude and present further issues raised by our analysis.

2. SOMALI VERBAL MORPHOLOGY

2.1 *General information*

Somali has a predominantly suffixal morphology. In particular, all inflectional markers (PNG (person, number, gender) and TAM (tense, aspect, mood)) are suffixes. Any verb form can be decomposed as follows:[2]

(1) *The Somali verb template* (adapted from Saeed 1993: 38–39)

(a)	(b)	(c)	(d)	(c′)
Root	Lexical morpheme (LEX)	Personal morpheme (PNG)	Tense-aspect-mood morpheme (TAM)	Personal morpheme (PNG)

As a representative example, consider the past and present tense paradigms of the verb *keːn* ‘bring’ in (2).

[2] In addition to regular (suffixed) verbs, there are five irregular verbs (Saeed 1999: 97–103). Four of them display prefixal morphology (*jiːl* ‘be (in a place)’, *jiri* ‘say’, *jimi* ‘come’ and *jiqiːn* ‘know’): the personal markers are prefixed to the verb stem and tense-aspect-mood categories are expressed by stem-internal vowel alternations (past *iː* vs. present *aː*). Finally, *ah* ‘be’ has an idiosyncratic pattern.

(2) keːn *'bring'*

	Past	Present
1S	keːn-aj[3]	keːn-aː
2S	keːn-t-aj	keːn-t-aː
3MS	keːn-aj	keːn-aː
3FS	keːn-t-aj	keːn-t-aː
1P	keːn-n-aj	keːn-n-aː
2P	keːn-t-eː-n	keːn-t-aː-n
3P	keːn-eː-n	keːn-aː-n

A form like *keːntaːn* (present, 2P) can be cast in template (1) as in (3).

(3) *Template for* keːntaːn *'bring PRES, 2P'*

(a)	(b)	(c)	(d)	(c′)
Root	LEX	PNG	TAM	PNG
keːn	Ø	t	aː	n
'bring'		2nd	present	plural

There are two TAM markers: *-aj-* (past) and *-aː-* (present). PNG markers are identical across tense categories, and appear in (4):[4]

(4) *PNG markers*

	SG	PL
1	Ø	n
2	t	t ... n
3M	Ø	Ø ... n
3F	t	

Finally, LEX (lexical morpheme) corresponds to markers that derive a verb from a verbal, nominal or adjectival stem, and convey various grammatical values. Some representative examples are given in (5):

(5)

LEX	V, N, ADJ	V
-am 'medio-passive'	mil 'dissolve s.t.'	milam 'dissolve'
-oːb 'inchoative'	ɖab 'truth'	ɖaboːb 'become true'
-tam 'reciprocal'	ul 'stick'	ultam 'fight each other with sticks'
-eː 'causative'	jar 'small'	jareː 'make small'
-o 'autobenefactive'	fur 'open'	furo 'open for oneself'

[2] In Somali orthography, the past tense suffix is transcribed either as *-ay* or *-ey*. Its phonetic realization depends on the quality of the stem vowel (Saeed 1993: 33). Since the exact phonetic properties of the past tense suffix are not relevant in this article, we adopt a unified transcription, *-aj*.

[4] PNG 2P and 3P are discontinuous markers. Person and gender features (1, 3 = Ø, 2 = *t*) are realized to the left of TAM. Number features (SG = Ø, PL = *n*) are realized to the right of TAM. The 1P marker *n* is realized to the left of TAM.

2.2 *Three classes*

Some LEX markers trigger complex alternations, affecting both the stem and the PNG markers. For this reason, Somali verbs are traditionally divided into three main morphological classes, depending on the identity of LEX and on its ability to trigger morphologically conditioned allomorphy:[5]

(6) *Correspondences between Classes et LEX*

Class 1	• No LEX • LEX that trigger predictable stem- and PNG-alternations, e.g. -tam, -am, -oːb
Class 2	LEX = causative Class 2a: LEX = -i Class 2b: LEX = -eː
Class 3	LEX = autobenefactive Class 3a: LEX = -so Class 3b: LEX = -o

As an example, consider the verb *mar* 'pass, tie up'. *Mar* can be suffixed by three distinct LEX elements: causative *-i* and autobenefactive[6] *-so* and *-o*, yielding the following three derived verbs: *mari* 'make pass, rub with' (class 2a), *marso* 'finish off, consume, dress up' (class 3a) and *maro* 'be finished/empty, become used up' (class 3b). The present tense paradigms of these verbs appear in (7). To complete the picture, we illustrate class 2b with the verb *ʕareː* 'anger', derived from the noun *ʕaro* 'anger'.

(7) *Basic conjugation (classes 1, 2a, 2b, 3a, 3b)*

		(a) Class 1	(b) Class 2a	(c) Class 2b
Imperative 2S		mar	mar-i	ʕar-eː
Infinitive		mar-i	mar-in	ʕar-ajn
Progressive (present)	1S	mar-aj-aː	mar-in-aj-aː	ʕar-ajn-aj-aː
Present	1S	mar-aː	mar-i-j(j)-aː	ʕar-eː-j(j)-aː
	2S	mar-t-aː	mar-i-s-aː	ʕar-aj-s-aː
	3MS	mar-aː	mar-i-j(j)-aː	ʕar-eː-j(j)-aː
	3FS	mar-t-aː	mar-i-s-aː	ʕar-aj-s-aː
	1P	mar-n-aː	mar-in-n-aː	ʕar-ajn-n-aː

[5] As mentioned in the introduction, there is a fourth class, which contains so-called hybrid verbs. This class exclusively contains adjectives, which are in fact stative verbs with either LEX = Ø or LEX = (*s*)*an*. Conjugation 4 behaves in a specific way, and for this reason, we will mention it only when needed for our reasoning.

[6] As noted by an anonymous *JL* referee, Saeed (1995) refers to this class as 'middle voice' verbs rather than autobenefactive verbs. For details on the semantics of these verbs, we refer to Saeed (1995).

	2P		mar-t-aː-n		mar-i-s-aː-n	ʕar-aj-s-aː-n
	3P		mar-aː-n		mar-i-j(j)-aː-n	ʕar-eː-j(j)-aː-n
			‘pass, tie up’		‘make pass’	‘anger’
		(d)	Class 3a	(e)	Class 3b	
Imperative 2S			mar-so		mar-o	
Infinitive			mar-san		mar-an	
Progressive (present)	1S		mar-san-aj-aː		mar-an-aj-aː	
Present	1S		mar-sad-aː		mar-t-aː	
	2S		mar-sa-t-aː		mar-a-t-aː	
	3MS		mar-sad-aː		mar-t-aː	
	3FS		mar-sa-t-aː		mar-a-t-aː	
	1P		mar-san-n-aː		mar-an-n-aː	
	2P		mar-sa-t-aː-n		mar-a-t-aː-n	
	3P		mar-sad-aː-n		mar-t-aː-n	
			‘finish off, dress up’		‘be finished, become used up’	

The matrix in (7) gives, for each class, the citation form (imperative second person singular (or: imperative 2S)), the infinitive, the progressive present[7] and the present tense paradigm. Past forms are parallel to present forms, the only difference being that they bear the past marker *-aj* instead of the present marker *-aː*. In particular, the allomorphic phenomena are identical in both tense paradigms. For this reason the data in (7) are sufficient for our purpose.[8]

2.3 *Allomorphic alternations*

Class 1 verbs differ from class 2 and 3 verbs in that they either have no LEX marker or bear a LEX marker that trigger predictable stem- and PNG-alternations. By contrast, class 2 and 3 verbs exhibit intriguing patterns of alternations that cannot at first sight be derived by the general phonological rules of the language.

2.3.1 *Deviant patterns in class 1 are triggered by general phonological rules*

In class 1, stem- and PNG-alternations derive exclusively from general phonological rules. We briefly review these rules below.

Consider first the verb *hadal* ‘talk’ (class 1) in (8a).

[7] The progressive present is traditionally analyzed as follows: infinitive + *aj*+ TAM, where *aj* is the remnant form of an auxiliary verb; see Saeed (1993: 43–44, 90, 93).

[8] We discuss the segmentation shown in (7) in Sections 3 and 4. Class 2a/b 1S, 3MS and 3P surface either with simple [j] or with [jj]. According to our informants, the two forms are in free variation.

(8) /l + t/ → [ʃ]

(a)	hadal	hadal-t-aː	→	hadaʃaː	‘you talk’
	talk. IMP 2S	talk-2S-PRES			
(b)	beːl	beːl-ta	→	beːʃa	‘the community’
	community	community-DET			
	abaːbul	abaːbul-to	→	abaːbuʃo	‘organizer.F’
	organize.IMP 2S	organize-F AGENTIVE			
	sagaːl	sagaːl-tan	→	sagaːʃan	‘ninety’
	nine	nine-unit of ten			

The 2S present form of this verb /hadal + t + aː/ surfaces as *hadaʃaː*, **hadaltaː*. This results from the application of the following rule: /l + t/ → [ʃ]. As illustrated in (8b), this rule applies in various morphological contexts and is known as a very general rule of the language (Saeed 1993: 26, 301).

The second source of phonological allomorphy in class 1 results from the application of a voicing rule affecting the dental plosive tː /t/ → [d][9] / V_V. For instance, the 2S present form of the verb *ʕi* ‘cry out’ (class 1), /ʕi-t-aː/, surfaces as [ʕidaː], as seen in (9a).

(9) /t/ → [d] / V_V

(a)	ʕi-t-aː	→	ʕidaː	‘you cry out’
	cry out-2S-PRES			
(b)	mindi-ta	→	mindida	‘the knife’
	knife-DET			
	hoːjo-ta	→	hoːjada	‘the mother’
	mother-DET			

Again, this rule applies in other contexts, too, for example, when the feminine article *-ta* is suffixed to a noun ending in a vowel, as in (9b) (Saeed 1993: 302).[10]

Third, class 1 CVCVC verb stems display vowel ∼ zero alternations across their present and past tense paradigms: they surface as CVCC- when followed by a vowel and as CVCVC elsewhere, as seen in (10).

(10) Ø → V_i / V_iC_C{C,#}

(a)	gudbaː	∼	gudubtaː		gudub	‘cross’
	PRES 1S		PRES 2S		IMP 2S	
(b)	nirgo	∼	nirigta		nirig	‘camel foal’
	N PL		N SG DET		N SG	
	afraːd	∼	afartan		afar	‘four’
	NUM ordinal		NUM unit of ten		NUM	
	‘fourth’		‘forty’		‘four’	

[9] Because of the lack of phonemic contrast between alveolar and dental plosives, we use a broad transcription of dental plosives: [d] instead of [d̪].

[10] We return to the *o* ∼ *a* alternation in *hoːjo* ∼ *hoːjada* in Section 4.3 below.

For instance, the second *u* of *gudub* 'cross' (class 1) appears in the citation form (*gudub* 2S imperative) and in forms involving a consonant-initial suffix (e.g. *gudubtaː* 2S present), only. The same vowel ∼ zero alternation occurs in the same context in nouns, as shown in (10b) (Saeed 1993: 27–28, 295).

The last two relevant phonological rules are neutralization rules. The first one targets final /t/ and /d/. The contrast /t/ vs. /d/ is neutralized in final position: both /t/ and /d/ surface as [d] in word-final position. For instance, /gunut/ 'tie a knot' (class 1) surfaces as *gunt-* when it is followed by a vocalic suffix (e.g. *guntaː* 1S present) but as *gunud* in final position, as in (11a), on a par with *hurud* 'go to sleep' (class 1) in (11b); see Saeed (1993: 30).

(11) /t/, /d/ → [d] / __#

		PRES 1S	IMP 2S	
(a)	/t/	guntaː	gunud	'tie a knot'
(b)	/d/	hurdaː	hurud	'go to sleep'

The second one targets /m/ and /n/. Both /m/ and /n/ surface as [n] in word-final position and before consonants. Hence *doːn* 'dredge' (class 1, 2S imperative), *doːntaː* (2S present) vs. *doːmaː* (1S present). The same situation obtains in nouns, as shown in (12b) (Saeed 1993: 301).

(12) /m/, /n/ → [n] / __ {C, #}

		/ __#	/ __C		/ __V	
(a)		IMP 2S	PRES 2S		PRES 1S	
	/m/	doːn	doːntaː	vs.	doːmaː	'dredge'
	/n/	doːn	doːntaː	vs.	doːnaː	'want'
(b)		SG	SG-DET		PL	
	/m/	tʃin	tʃinta	vs.	tʃimo	'arm'
	/n/	dan	danta	vs.	dano	'aim'

Since the application of the rules above is fully predictable on the basis of the phonological context, the verbs of the type mentioned in this section are correctly analyzed as belonging to the same class – class 1.

2.3.2 *Deviant patterns in classes 2 and 3 cannot be derived from general phonological rules*

In classes 2 and 3, both LEX and PNG morphemes display allomorphic alternations that cannot be easily derived from general phonological rules. The table in (13) summarizes the identity of the PNG markers involved in class 1, 2a and 2b, and 3a and 3b. PNG markers are identical within a given class: class 2a and 2b on the one hand, class 3a and 3b on the other hand, have the same PNG markers. The shaded cells contain the forms that are unexpected.

(13) *PNG allomorphy*

PNG	Class 1	Class 2a/b	Class 3a/b
1S	Ø	Ø	Ø
2S	t	s	t
3MS	Ø	Ø	Ø
3FS	t	s	t
1P	n	n	n
2P	t	s	t
3P	Ø	Ø	Ø

The PNG markers involved in class 2a/b differ from those of class 1 in the following respect:

(14) 2S/3FS/2P /t/ surfaces as [s], e.g. 2S present *mari-s-aː* 'you make pass'.
Given rule (9), we would expect **mari-d-aː*.

At first sight, the png markers involved in class 3a/b are identical to those of class 1. This apparent identity, however, hides an intriguing fact:

(15) 2S/3FS/2P /t/ surfaces as [t] in intervocalic positions, e.g. 2S present *mar-sa-t-aː* 'you finish off' (class 3a), *mar-a-t-aː* 'you are finished, become used up' (class 3b).
Given rule (9), we would expect **mar-sa-d-aː*, **mar-a-d-aː*.

We now turn to the allomorphy of LEX. Class 1 verbs display no unexpected allomorphy. Thus, the table in (16) represents the neutral case. The shaded cells in (16) as well in all remaining tables correspond to forms that do not exist.

(16) *Class 1: No allomoprphy*

	Stem	LEX	PNG	TAM
Imperative 2S	mar	Ø		
1S/3MS/3P (= before V)			Ø	aː
2S/3FS/2P (= before C except *n*)			t	
1P (= before *n*)			n	

Classes 2a/b and 3a/b involve various allomorphs, whose segmentation is not straightforward.

Following the traditional view (presented in (6) above), we assume that LEX = *i* in class 2a. Then, we isolate the PNG and TAM markers and arrive at the table in (17). The column headed '?' (between LEX and PNG) contains the material that is left unattributed.

(17) *Class 2a allomorphy*

	Stem	LEX	**?**	PNG	TAM
Imperative 2S	mar	i			
1S/3MS/3P (= before V)			**j(j)**	Ø	aː
2S/3FS/2P (= before C except *n*)				s	
1P (= before *n*)			**n**	n	

In class 2b, LEX has two allomorphs: *-eː-* and *-aj-*, yielding the following segmentation:

(18) *Class 2b allomorphy*

	Stem	LEX	**?**	PNG	TAM
Imperative 2S	mar	eː			
1S/3MS/3P (= before V)			**j(j)**	Ø	aː
2S/3FS/2P (= before C except *n*)		aj		s	
1P (= before *n*)			**n**	n	

The appearance of *-j(j)-* under '?' in (17) and (18) could be considered a phonotactically conditioned *j*-insertion (hiatus resolution). Along these lines, we should not posit an allomorphy in 1S/3MS/3P. However, this cannot be right. There is indeed a clear contrast between *i*-final class 1 verbs, e.g. *bari* 'spend the night' and class 2a verbs, e.g. *mar-i*. Class 1 1S /bari-Ø-aː/ surfaces as [barjaː]. By contrast, class 2 1S /mar-i-Ø-aː/ may surface as [marjaː], [marijaː] and [marijjaː]. Crucially only class 2 1S may be realized as [marij(j)aː]. We take this fact to indicate that the underlying structure of class 2 1S is not parallel to that of class 1 1S: it involves allomorphy.

(19)

Class 1, 1S			Class 2a, 1S		
/bari-Ø-aː/	→	[barjaː]	/mar-i-j-aː/	→	[marjaː]
		*[barijaː]			[marijaː], [marijjaː]

Turning to class 3, we present in (20) and (21) the distribution of the allomorphs of LEX as well as the material left unidentified under '?'.

(20) *Class 3a allomorphy*

	Stem	LEX	**?**	PNG	TAM
Imperative 2S	mar	so			
1S/3MS/3P (= before V)		sa	**d**	Ø	aː
2S/3FS/2P (= before C except *n*)				t	
1P (= before *n*)			**n**	n	

(21) *Class 3b allomorphy*

	Stem	LEX	**?**	PNG	TAM
Imperative 2S	mar	o			
1S/3MS/3P (= before V)		Ø	**t**	Ø	aː
2S/3FS/2P (= before C except *n*)		a		t	
1P (= before *n*)			**n**	n	

The tables in (17), (18), (20) and (21) set the terms of the debate: what is the relation between the material appearing under '?' and the allomorphs of LEX? We will take a position on this question in the following sections.

3. THE ALLOMORPHY OF THE CAUSATIVE MARKER

We first focus on class 2a (causative *-i*) and class 2b (causative *-eː*) and propose a unique representation of the causative suffix in both classes, thus unifying classes 2a and 2b under a single inflectional class. We further show that the alternations observed on the surface result from the application of regular phonological principles to this underlying representation. Our line of reasoning necessitates first a detour into the status of *-s* in class 3a, in Section 3.1. Sections 3.2 and 3.3 are dedicated to the analysis of *-i* and *-eː*, respectively.[11]

3.1 *The identity of -s in class 3a: Allomorphy of the causative suffix in derivational morphology*

A comparison of the tables in (20) and (21) above shows that class 3a and class 3b basically differ in the presence vs. the absence of *-s-* in LEX.[12] In order to gain insight into the status of this segment, let us consider again the three verbs derived from *mar* 'pass, tie up' in (22a): *mari* 'make pass' (causative, (22b)), *maro* 'be finished/empy, become used up' (autobenefactive, (22c)), and *marso* 'finish off, consume, dress up' in (22d):

(22) *Comparison between* mar, mari, maro *and* marso

(a)	mar					'pass, tie up'
(b)	mar-i	=	mar	-	i	'make pass'
			'pass'		CAUS	
(c)	mar-o	=	mar	-	o	'be finished, become used up'
			'tie up'		AUTOBEN	
(d)	mar-so	=	?			'finish off, dress up'

Class 2a verbs (e.g. *mari*) are derived from class 1 verbs by suffixation of *-i*, and they are causatives (23b). Class 3b verbs (e.g. *maro*) are derived from class 1 verbs by suffixation of *-o*, and they are autobenefactives (23c). Class 3a verbs (e.g. *marso*) are derived from class 1 verbs by suffixation of *-so* and they have both an autobenefactive and a causative meaning (23d). The relationship between *maro* and *mar* is identical to that between *marso* and *mari*: *maro* is the autobenefactive counterpart of *mar*, *marso* is the autobenefactive counterpart of *mari*. We conclude that *-s* in class 3a is an allomorph of the causative: autobenefactive verbs in *-so* are derived from causative verbs by 'replacing' *-i* by *-s* and adding autobenefactive *-o*:

[11] The analysis builds on previous work by Barillot & Bendjaballah (1998).

[12] For the moment, we abstract away from other contrasts (e.g. *sad* (CONJ 3a) vs. *t* (CONJ 3b)). We analyze the autobenefactive suffix in Section 4.

(23) *The segmentation of* marso *'finish off, consume, dress up'*

(a)	mar							'pass, tie up'
(b)	mari	=	mar	+	i			'make pass'
			'pass'		CAUS			
(c)	maro	=	mar	+			o	'be finished'
			'tie up'				AUTOBEN	
(d)	marso	=	mar	+	s	+	o	'finish off, consume'
			'pass'		CAUS		AUTOBEN	

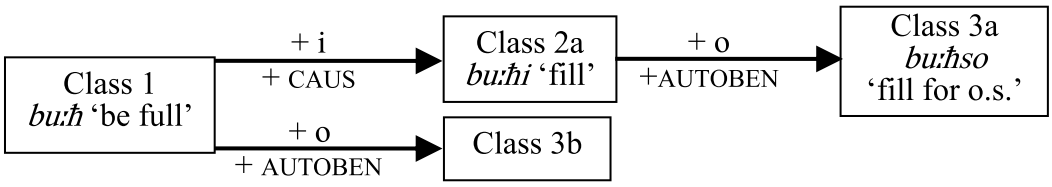
The derivation of class 3a verbs on the basis of class 2a verbs is very productive; a representative set of examples is given in (24).

(24) *Class 3a verbs deriving from class 2a verbs*

Class 1	Class 2a	Class 3a
buːħ	buːħi	buːħso
'be full'	'fill'	'fill for o.s.'
ɖeːf	ɖeːfi	ɖeːfso
'benefit'	'nourish, benefit s.o.'	'benefit from, nourish o.s.'
ɖoːf	ɖoːfi	ɖoːfso
'leave'	'make s.o. leave, export'	'export'
did	didi	didso
'disperse'	'disperse s.t.'	'chase away for o.s.'
dillaːʕ	dillaːʕI	dillaːʕso
'be cracked'	'crack, tear s.t. off'	'tear, crack for o.s.'
fid	fidi	fidso
'spread (INTR)'	'spread s.t. out'	'spread for o.s.'
hub	hubi	hubso
'be sure'	'make sure'	'make sure for o.s.'
tʃab	tʃabi	tʃabso
'get broken'	'break s.t.'	'break s.t. for o.s.'
kar	kari	karso
'be boiling'	'boil, cook s.t.'	'cook for o.s.'
qoton	qotomi	qotonso
'be vertical'	'put s.t. vertical, erect'	'erect for o.s.'
soːf	soːfi	soːfso
'go out to graze'	'drive (livestock) out to graze'	'take (livestock) to pasture'
tir	tiri	tirso
'cancel'	'count'	'count for o.s.'
urur	ururi	ururso
'gather (INTR)'	'gather, collect'	'gather, collect for o.s.'
jeːr	jeːri	jeːrso
'call out, say'	'dictate, cause to say'	'take dictation for o.s.'

The derivational paths involving class 1, 2a, 3a/b verbs are summarized in (25):

(25) *Derivational paths*



An additional argument in favor of the analysis of *-s* as an allomorph of the causative comes from the behavior of CVCVC verbs, e.g. *gudub* 'cross'. As mentioned above, in (10), these verbs display a vowel ~ zero alternation conditioned by the context: imperative 2S *gudub* ~ present 1S *gudbaː*. Following Barillot (2002), we assume the stem of such verbs to be /$C_1VC_2C_3$/. The stem vowel propagates onto the position between the second and the third stem-consonant when the phonotactic constraints of the language require this position to be identified. This is the case in two configurations: (i) if the verb stem is in final position (i.e. 2S imperative), e.g. *gudub* ~ **gudb*, CC# clusters are prohibited in Somali; and (ii) if the verb stem is followed by a C-initial suffix (i.e. 2S/3FS/1P/2P), e.g. *gudubtaː* ~ **gudbtaː*, CCC clusters are prohibited in Somali.[13] If the stem is followed by a V-initial suffix, no vowel needs to surface between the second and the third stem-consonants, e.g. PRES 1S *gudbaː*.

Given this distribution, we expect a class 2a causative derived from a CVCVC verb to have the following shape: CVCC*i*. The causative derived from *gudub* 'cross' is indeed *gudbi* 'send across'. Consider now the autobenefactive derived from a CVCC*i* causative. We expect *-so* to attach to the stem /gudb/: **gudbso*. The ban on CCC clusters prevents this form from surfacing. Two equally possible strategies are conceivable:

(26) *Autobenefactive of CVCC$_i$ verbs: two strategies*

(a) Propagation of the stem vowel between the last two stem-consonants, yielding CV_iCV_iC*so*.

(b) Realization of a vowel between the final stem-consonant and *-so*, yielding CV_iCCV_j*so*.

There are 57 class 3a verbs derived from a CVCVC stem (see Agostini, Puglielli & Siyaad 1985 and Zorc 1993). Fifty-two of them are of the type CV_iCV_iC*so*, seen in (27a) below. The remaining forms are distributed as follows: four are CVCC*iso* verbs, illustrated in (27b), and there is one verb with both CV_iCV_iC*so* and CVCC*iso*, with the same meaning, seen in (27c). We take the sequence *-is-* in (27b–c) to be an allomorph of the causative. This allomorph is selected to prevent a CCC cluster from surfacing.

[13] As an anonymous *JL* referee suggests, this can alternatively be expressed as follows: 'No homosyllabic CC clusters are allowed'.

(27)

	Class 1	Class 2a	Class 3a
(a)	qurun	qurmi	qurunso
	‘rot, stink’	‘cause to rot, stink’	‘cause to rot, stink for o.s.’
	taran	tarmi	taranso
	‘be multiplied’	‘cause to multiply’	‘get more for o.s.’
(b)	ereg	ergi	ergiso
	‘give s.t. on trust’	‘lend s.t. to s.o.’	‘take s.t. in trust’
(c)	koroɖ	korɖi	korɖiso, koro̲ɖso
	‘be increased’	‘cause to grow, increase’	‘increase for o.s.’

A similar situation obtains in the derivation of stative verbs. Consider as an example *marsan* and *maran* in (28a): *maran* ‘be tied up’ is semantically related to *mar* ‘tie up’, whereas *marsan* ‘be rubbed with s.t.’ is built on *mari* ‘rub s.o. with s.t.’. Since *-an* is the stative suffix, the *-s-* of *marsan* is an allomorph of the causative suffix. The elements *-s-* in *marsan* and *-s-* in *marso* are one and the same allomorph of the causative suffix. In the derivation of statives from CVCC stems, the strategies illustrated in (27) above obtain, giving CV_iCV_iC*san* in (28b), CVCC*isan* in (28c), and in one case, again, both strategies result in two forms with the same meaning, as shown in (28d).

(28)

	Class 1	Class 2a	Class 4
(a)	mar	mari	marsan
	‘pass, tie up’	‘make pass, rub s.o. with s.t.’	‘be rubbed with s.t.’
			maran
			‘be tied up’
(b)	debeʕ	debʕi	debe̲ʕsan
	‘become loose’	‘loosen’	‘be loose’
	damaʕ	damʕi	dama̲ʕsan
	‘desire s.t.’	‘make s.o. desire s.t.’	‘be desirous of s.t.’
(c)	ereg	ergi	ergi̲san
	‘give s.t. on trust’	‘lend s.t. to s.o.’	‘be lend’
(d)	firiɖ	firɖi	firɖi̲san, firi̲ɖsan
	‘scatter (INTR)’	‘scatter (TR)’	‘be scattered’

To conclude, the causative suffix has three allomorphs, *-i*, *-s-* and *-is-*,[14] whose distribution is as follows:

[14] Two types of additional evidence that *-(i)s-* is an allomorph of the causative marker may be mentioned. First, consider *bari* ‘pass the night in peace’, an *i*-final class 1 verb in (i) below. Class 2a causative is derived by suffixation of *-i*, yielding *bari:*. The corresponding autobenefactive is *bari:so*. The causative marker must be *-is-*.

(i)

Class 1	Class 2a (*-i*)	Class 3a (*-so*)
bari	bari:	bari:so
‘pass the night in peace’	‘cause to pass the night in peace’	‘stay at dawn where we slept’
—	wejddi:	wejddi:so
	‘ask s.o. s.t.’	‘ask s.o. s.t. for o.s.’

(29) *Allomorphy of the causative suffix*

(a)	In isolation	*-i*	buːħi ‘fill’
(b)	Before a derivational suffix		
	i. VC __	*-s-*	buːħso ‘fill for o.s.’
	ii. {CC, V} __	*-is-*	ergiso ‘take s.t. in trust’

We now have a complete description of the causative allomorphy and are in a position to propose a representation for this marker.

3.2 *The representation of the causative marker*

In this section, we argue that the underlying representation of the causative marker is /It/. First, we introduce Element Theory (Kaye et al. 1985, 1990), which we exploit to decompose the vocalic segments, and the CV-framework (Lowenstamm 1996), which we adopt to represent the skeletal tier of phonological representations. Then we propose our representation of the causative marker. Finally, we focus on three inflected forms, 1S, 2S and 1P, which illustrate how the underlying material is computed into surface forms by the phonology of the language.

3.2.1 *Element Theory and the representation of the causative marker*

Element Theory (Kaye et al. 1985, 1990) is a theory of segmental representations. In this framework, segments are the surface realizations of underlying Elements, rather than features. Elements can be combined by an operation called ‘Fusion’. Fusion is an asymmetrical operation, involving a head (underlined in the representations) and an operator (Kaye et al. 1985: 309). Following Backley (2011), we assume the Elements associated with vowel structure to be |A|, |I| and |U|. The internal structure of consonants will not be relevant in this article; we thus simply posit the segment, as a shortcut to a more complex underlying Element structure. Abstracting away from harmony phenomena, which are irrelevant for

---

The second additional piece of evidence comes from a comparison of the argument structure of the verbs ending in *-so* (class 3a) and *-o* (class 3b). An exhaustive examination of Agostini et al. (1985) reveals that only 3% of 550 verbs in *-o* add an argument to the base while 44% of 500 verbs in *-so* increase the number of arguments of the base (see Barillot & Bendjaballah 1998). The following are representative examples:

(ii)

Class 1	Class 3b (*-o*)	Class 3a (*-so*)
adeːg	adeːgo	adeːgso
‘serve (INTR)’	‘do tasks for oneself (INTR)’	‘employ somebody (TR)’
aruːr		aruːrso
‘be gathered (INTR)’		‘gather for oneself (TR)’
baːr	baːro	
‘inspect (TR)’	‘inspect for oneself (TR)’	
daːq		daːqso
‘graze, eat (grass) (TR)’		‘feed (livestock) with s.t. (DITR)’

our purposes, we derive the Somali underlying five-vowel system as in (30); see Backley (2011: 41).

(30) *Fusion*

[a] = |A|
[i] = |I|
[u] = |U|
[e] = |A I|
[o] = |A U|

As for the skeletal level of phonological representations, we adopt the CV-framework (Lowenstamm 1996, Scheer 2004), according to which the skeletal tier consists of a strict alternation of non-branching nuclei (C) and non-branching onsets (V). In this framework, the distribution of empty V positions is constrained by the Empty Category Principle and Proper Government (PG), as defined in Kaye et al. (1990: 219) and subsequent work, e.g. Charette (1990: 236):

(31) *Empty Category Principle*
A position may be uninterpreted phonetically if it is properly governed.

(32) *Proper Government*
A properly governs B iff
(i) A governs B (A and B are adjacent on the nuclear projection);
(ii) A is not licensed;
(iii) No governing domain intervenes between A and B.

The notion of licensing has been discussed in various contexts (Charette 1989, 1991; for details within the CV-framework, see Scheer & Ségéral 2001a: 138). Licensing determines whether a segment may be realized or not: typically, an unlicensed segment has a very restricted phonetic expression, or none at all. Licensing was introduced to encode the observation that there is a dependency between the onset and the nucleus: the segmental expression of an onset crucially depends on the ability of its nucleus to license. More specifically, the ability of a nucleus to license is constrained as follows:

(33) *Licensing*

An empty nucleus may neither govern nor license.

A filled nucleus may both govern and license.

Finally, following Bendjaballah & Haiden (2008) and Lowenstamm (2008), we assume that skeletal units can be the exponents of grammatical features. The minimal skeletal unit being CV, this means that a CV unit can be a morpheme.

Equipped with these tools, we argue that the Somali causative marker has the following representation:

(34) CAUS CV

I t

This structure contains (i) the minimal skeletal unit, CV, and (ii) two objects at the segmental level: |I| and *t*. As can be seen from (34), we assume that |I| and *t* are floating, i.e. they are not lexically linked to the skeleton. The question is thus whether their association to the skeleton is constrained by some principles, and if it is, by which ones. We assume that the association of |I| is constrained by the Empty Category Principle and Proper Government. Let us now consider *t*. The table in (35) gives a survey of the distribution of [t] in Somali.

(35) *Distribution of [t]*

			Agostini et al. 1985	Zorc 1993	Keenadiid 1976
(a)	Initial	#tV	1512	1041	584
	Onset after coda	... CtV ...	1280	960	481
	Intervocalically	... VtV	1026	690	370
(b)	Final	t#	1 $_{(L)}$	14 $_{(L)}$	1 $_{(L)}$
	Coda	... tC ...	11 $_{(L)}$	22 $_{(L)}$	2 $_{(L)}$
	Geminate	... tt ...	0	1 $_{(L)}$	0

The table is based on an exhaustive examination of three of the most comprehensive Somali dictionaries: Agostini et al. (1985; Somali–Italian, 36,276 items), Zorc (1993; Somali–English, 25,987 items) and Keenadiid (1976; monolingual Somali, 14,497 items). It turns out that [t] is almost always followed by a vowel, as seen in (35a, b). [t] appears before a consonant or in final position in a small number of words, only (40 out of 37,000 words).[15] Moreover, all of them are loanwords ($_L$ in (35b)), e.g. *matħaf* 'museum' < Arabic *mutħaf*, *batro:l* 'petrol' < English.

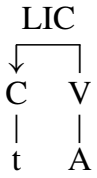
We thus assume that [t] must be followed by a vowel in Somali. We take this distributional observation to indicate that, in order to be linked to a C position, *t* must be followed by a phonetically interpreted V position.[16] In our framework, this generalization is formulated in (36):[17]

[15] Among the 40 words in which [t] appears either in final position or after a consonant, 37 items are found in Zorc (1993) and three items in Agostini et al. (1985), which are absent from Zorc (1993).

[16] For a similar constraint on glides in Berber, see Guerssel (1990: 44–47).

[17] In addition to licensing by following V, a special case of licensing will be mentioned in the article: licensing by virtue of being part of a geminate structure. On geminates as governing domains, see e.g. Guerssel (2003).

(36) t *must be licensed*



We now review the different configurations involving CAUS and show how the surface forms are derived.

3.2.2 *Causative allomorphy (I): Derivational morphology*

The causative marker appears either in isolation or before the autobenefactive marker *-o*. In isolation, CAUS always surfaces as *-i*:

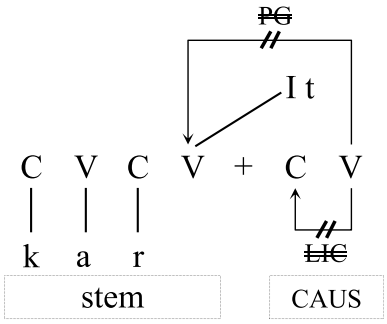
(37)

	Base		Causative	
(a)	CVC	kar ‘boil’ [intr]	CVC*i*	kari ‘boil s.t., cook’
(b)	CVVC	buːħ ‘be full’	CVVC*i*	buːħi ‘fill’
(c)	CVCVC	ereg ‘entrust s.t.’	CVCC*i*	ergi ‘entrust s.t. to s.o.’

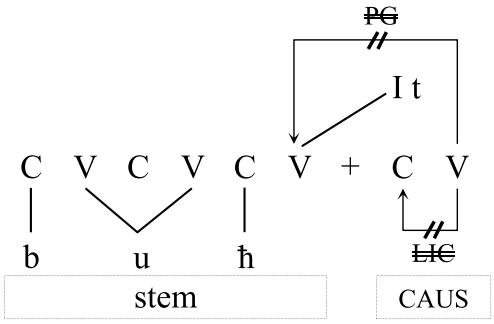
The three configurations exemplified in (37) are illustrated in (38).

(38) *CAUS in isolation*

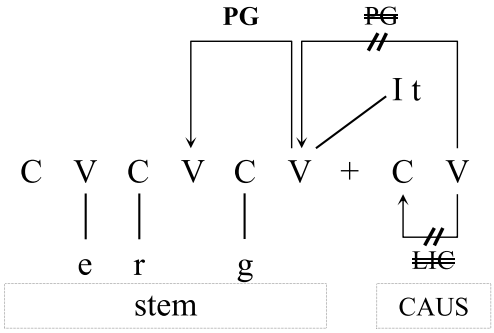
(a) *kari* ‘boil s.t.’



(b) *buːħi* ‘fill’



(c) *ergi* ‘entrust s.t. to s.o’



In all cases, the final V position of the verb stem is not properly governed, and it is identified by caus |I|. The segment *t* cannot be associated with the skeleton: it is not followed by a vowel, and is not licensed.

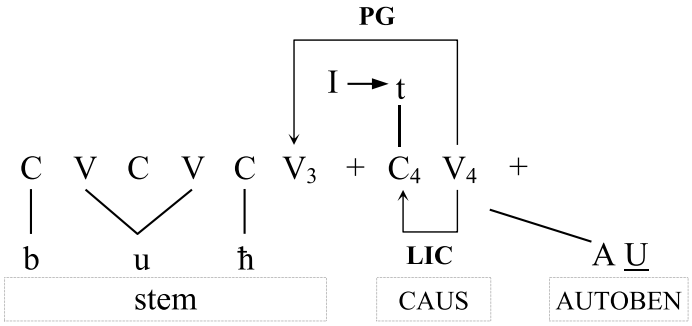
If CAUS is followed by the autobenefactive marker, it surfaces either as *-s-*, as in (39a), or as *-is-*, as in (39b):

(39)

	Base		CAUS	CAUS + AUTOBEN
(a)	CVC	mar ‘pass’	mari ‘make pass’	marso ‘finish off, dress up’
	CVVC	buːħ ‘be full’	buːħi ‘fill’	buːħso ‘fill for o.s.’
(b)	CVCVC	ereg	ergi	ergiso
		‘entrust s.t.’	‘entrust s.t. to s.o.’	‘take s.t. in trust’

We start with the derivation of CAUS + AUTOBEN from a CV(V)C base in (39a). As can be seen in (40), *-o* (resulting from A U) identifies the final V position.[18]

(40) *Causative + autobenefactive, CV(V)C verbs:* buːħ-s-o *‘fill for o.s.’*

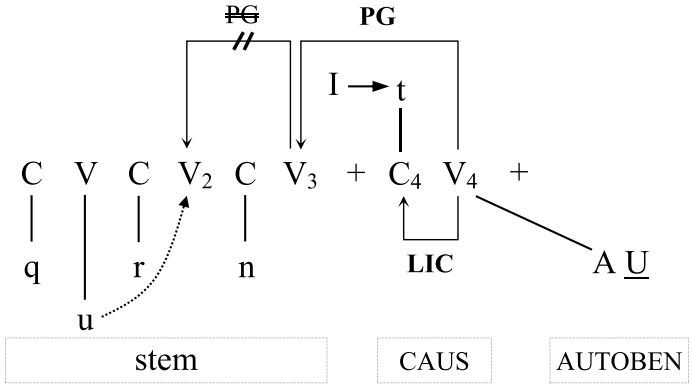


This has two consequences: (i) this position licenses *t* to be associated with the preceding C position, and (ii) it governs the penultimate V position (V_3) with the effect that |I| does not need to be associated with that position. We claim that *t* is palatalized to [s] by underlying |I|, and caus surfaces as [s].

Consider now the derivation of CAUS + AUTOBEN from a CVCVC base in (39b) above, as illustrated in (41).

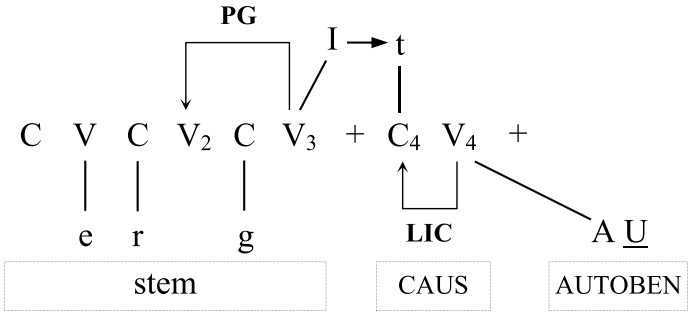
(41) *Causative + autobenefactive, CVCVC verbs*

(a) *qurunso* ‘cause to rot’



[18] We turn to the exact representation of AUTOBEN in Section 4.1 below.

(b) *ergiso* 'take s.t. in trust'



As in (40), *-o* on the final V position (V_4) licenses *t* to be associated with C_4. It also governs V_3. In this configuration, two options are available: either V_3, being governed, remains empty, and V_2 must be identified, yielding *qurunso*, in (41a), or |I| is associated with V_3, and V_2 remains empty, which yields *ergiso*, in (41b). The former option consists in the propagation of the stem vowel between the last two stem-consonants and is attested in 52 cases (e.g. *qurunso* 'cause to rot', (26a) and (27a) above). The latter option is attested in only four verbs (e.g. *ergiso* 'take s.t. in trust'). Finally, one verb displays both strategies: *koroɖso*, *korɖiso* 'increase for o.s.'.[19]

Before turning to the allomorphy of CAUS in inflectional morphology, we believe that it is important to reaffirm our hypothesis on how underlying /t/ is palatalized in (41b). An anonymous *JL* referee points out that, since [it] sequences are attested in Somali, |I| cannot both palatalize and be associated with the skeleton. This would imply that only floating |I| can palatalize and (41b) is not well-formed. We do not agree with this conclusion. First, note that the strategy presented in (41b) involves four verbs only. Secondly, an exhaustive count of the lexical entries given in two reference dictionaries (Agostini et al. 1985, Zorc 1993) reveals that there are 108 occurrences of [it] in Somali. These 108 occurrences can be divided into three groups:

(i) In 57 cases, *-i-* and *-t-* are heteromorphemic: 29 sequences involve the concatenation of *-i* and the derivational suffix *-taːn*, e.g. *fur* 'open' vs. *furitaːn* '(act of) opening'; 15 sequences result from several derivational processes, e.g. *baɦ* 'exit, leave, go out', *biɦi* 'take out, remove' and *biɦitin* 'leave, make a trip'; and 13 sequences correspond to compounds, e.g. *rabbituːg* 'prayer to God' < *rabbi* 'Master, the Lord' + *tuːg* 'pray'.

(ii) In 32 cases, the forms are clearly loanwords, e.g. *kitaːb* 'book' < Arabic *kitaːb*, (ʕ)*isbitaːl* 'hospital' < English *hospital*.

(iii) Finally, there are 19 items in which the sequence *-it-* does not seem to be analyzable in terms of morphological structure, e.g. *abitej* 'kite', *fiːto* 'pimple, wound' and *hitiq* 'walk slowly'.

[19] Note that the theory does not predict which strategy is more likely to be attested: both are equally possible. Somali enforces propagation of the stem vowel in 53 cases out of 57. In fact, as extensively analyzed in Barillot (2002), this is the regular strategy in the language.

In (i), a morphological boundary intervenes between *-i* and *-t*. Such is not the case in the causative suffix. We assume that palatalization is blocked by the presence of a morphological boundary in group (i). As for group (ii), it is well-known that, under certain conditions, loanwords resist full integration into the native language phonology (see, among many others, Kang 2011: 2261, and references therein). Finally, we are left with 19 monomorphemic occurrences of [it], for which we have no explanation for the absence of palatalization. Such a situation is not uncommon. Consider Italian, for example. In Italian, palatalization is triggered by the plural suffix *-i* and affects velars, as in e.g. [amiːko] 'friend.M.SG' vs. [amiːtʃi] 'friend.MPL'. However, in certain comparable structures, palatalization does not apply, e.g. [pikko] 'peak.M.SG' and [pikki] 'peak.M.PL'. In Italian as in Somali, a regular phonological process (palatalization) may not apply under certain (lexical) conditions. In Somali, one possibility would be to assume that this condition is the presence of inherent skeletal material. The causative suffix introduces its own templatic space, the CV unit in (34) above (Bendjaballah 1998, 1999: 191ff.). In (40), as in (41), |I| associates with /t/, yielding /s/. In both configurations, C_4 is licensed by *-o* in V_4. As a consequence /s/ is linked to C_4. In (40) and (41a), |I| remains floating. In (41b), it is associated with V_3 and surfaces as [i]. The difference between (40) and (41a) on the one hand and (41b) on the other hand is that in (41b) [i] must surface because its absence would result in an ungrammatical CCC cluster. Both configurations are consistent with previous analyses of palatalization in Government Phonology and CVCV (see e.g. among many others, Charette 1989, Scheer & Ségéral 2001b, Cristófaro Silva 2003). The same referee seems to suggest that only floating |I| may palatalize. If we followed this suggestion, we would have to claim that *-i-* in *ergiso* is an epenthetic vowel (i.e. it would be similar to [i] in loanwords: *ʃifti* < shift (English), *waqti* < *waqt* 'time' (Arabic)). However, it is obviously not the case that in Somali only floating |I| may palatalize. Consider for instance the pairs *noːg* 'be tired' ∼ *noːtʃi* 'tire' and *daːq* 'graze' ∼ *daːtʃi* 'make graze'. In all comparable contexts, both palatalization of {g, q} to [tʃ] and realization of causative *-i* are observed (see Bendjaballah 1998, 1999 for analysis). We therefore maintain our analysis of *-i-* in *ergiso* as the causative *-i*.

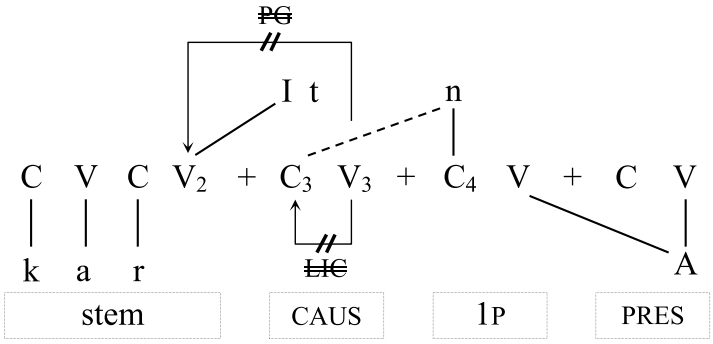
3.2.3 *Causative allomorphy (II): Inflectional morphology*

Recall from Section 2.3.2 above that allomorphy involves 1S/3MS/3P *-j(j)-* and 1P *-n-* in the column labeled '?' in the tables in (17) and (18). We select the following representative forms 1S, 2S and 1P and show that *-j(j)-* and *-n-* result from the application of regular phonology in the linearization process.[20]

We start with 1P *karinnaː* 'we cook s.t.' in (42).

[20] By 'linearization process', we mean the stage of derivation during which phonology applies and creates surface forms.

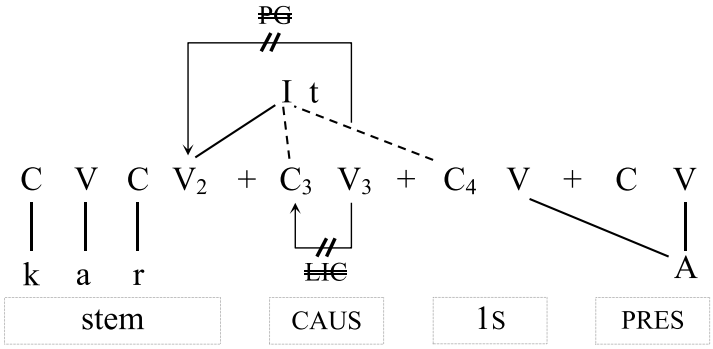
(42) *Class 2a, 1P, e.g.* karinna:*'we cook s.t.'*



The stem *kar* is followed by CAUS /It/, followed by PNG *n* and TAM |A|. We assume each of the two categories PNG and TAM to be equipped with a single CV unit (see the typology of spell-out proposed by Bendjaballah & Haiden 2008). The V position of CAUS, V_3, is empty. As a consequence (i) *t* cannot associate with its docking site C_3, and (ii) since V_2 is not properly governed, |I| associates with V_2. We are left with an empty C_3V_3 sequence, which is identified by the propagation of 1P *n* to C_3.

Consider now 1S *karij(j)a:* 'I cook s.t.' in (43).

(43) *Class 2a, 1S, e.g.* karij(j)a: *'I cook s.t.'*

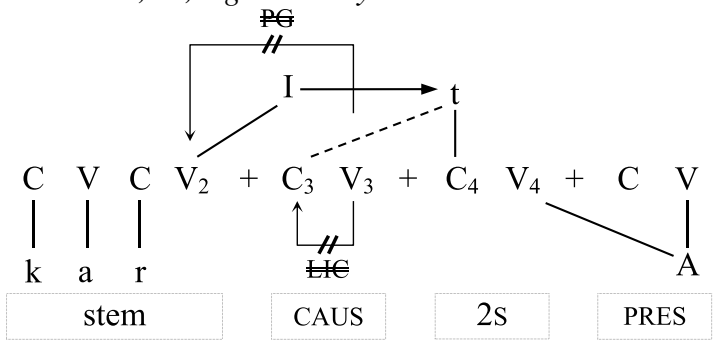


As in 1P, the V position of CAUS, V_3, is empty. As a consequence (i) *t* cannot associate with its docking site C_3, and (ii) since V_2 is not properly governed, |I| associates with V_2. We are left with an empty $C_3V_3C_4$ sequence. The only strategy making it possible to avoid such a sequence is the propagation of |I| to C_3 and C_4, hence [jj].[21]

Finally consider 2S *karisa:* 'you cook s.t.' in (44).

[21] A certain degree of variation exists among speakers: both [j] and [jj] are possible. This variability in the phonetic implementation of /jj/ does not affect the fact that, phonologically, it is geminated; see the discussion at the end of this section.

(44) *Class 2a, 2S, e.g.* karisaː *'you cook s.t.'*



V_3 being empty again, the same consequences obtain: (i) since the causative *t* is not licensed, it cannot associate with the skeleton, and (ii) since V_2 is not properly governed, |I| associates with V_2. The crucial difference with (43) is that PNG is not zero, but *t*. This *t* associates with C_4 (which is licensed by |A| in V_4), and propagates onto C_3. In other words, we have an underlying /tt/. /tt/ is palatalized by |I|, and surfaces as [s]. Note that [s] is a latent geminate in the sense that it is phonologically geminated, but phonetically realized as a single consonant (Scheer & Ségéral 2001a, 2001b).[22] Q4

The forms 1S, 2S and 1P are parallel in the sense that they involve an underlying geminate. The difference lies in the realization of this underlying geminate: it is always phonetically realized as a geminate in 1P ([karinnaː], *[karinaː]), it is optionally realized as a geminate in 1S ([karijjaː], [karijaː]) and it is never realized as a geminate in 2S ([karisaː], *[karissaː]). This result is coherent with the general behavior of *n*, *j* and *s* in Somali: [n] is always phonologically short, whereas [j] and [s] can be either short or long.

The evidence comes from a wide range of morphophonological processes in Somali: the pattern of V ∼ Ø alternations in CVCVC verbs, reduplication, compounding, behaviour of *n* + C clusters, loanwords, etc. (see Barillot 2002: 197–351 for a precise review and analysis). In this article we exemplify the reasoning on the basis of the behaviour of CVCVC verbs, e.g. *gudub* 'cross' in (45a).

[22] An anonymous *JL* referee suggests that |I| in (44) could not palatalize PNG *t* because it is not adjacent to it. Our analysis follows standard principles like the Obligatory Contour Principle (OCP, Leben 1973), according to which two identical phonological objects cannot be adjacent at the same level of representation. For this reason, the two *t*s in (44) (CAUS *t* and PNG *t*) must either fuse, or one of them must be discarded. In both cases, |I| is locally adjacent to PNG *t* (that is associated with C_3 and C_4) and it palatalizes *t* to *s*. In addition, note that the propagation of *t* to C_3 in (44) takes place even if the (empty) V_3 position cannot licence *t* to be associated with C_3. This is because geminates constitute a governing domain: PNG *t* is linked to C_4 (V_4 is not empty), and it licenses *t* to be associated with C_3, as the branch of a geminate structure. (On this property of geminates, see standard work in Government Phonology, e.g. Guerssel 2003.)

(45)

	IMP 2S	PRES 1S		
(a)	gudub	*gudu̲baː	gudbaː	‘cross’
(b)	feker	fek**e̲**raː	*fekraː	‘think’
	/fekker/	/fekkeraː/	*/fekkraː/	
(c)	beddel	bedde̲laː	*beddlaː	‘change’

Recall from Section 3.1 that these verbs are /CVCC/ at the phonological level. Some CVCVC verbs, however, resist V ∼ Ø alternation, e.g. *feker* ‘think’ in (45b). Building on the fact that CVC_iC_iVC verbs, e.g. *beddel* ‘change’ in (45c) never exhibit V ∼ Ø alternation, Barillot (2002) argues that intervocalic *k* in *feker* is phonologically long: /fekker/. The medial geminate forces the lexical vowel to propagate. If it did not propagate, a banned CCC cluster would surface.[23]

Examining the behavior of the complete set of CVC_0VC verbs in Agostini et al. (1985) and Zorc (1993), Barillot (2002) shows that the phonological status of Somali consonants is as summarized in the table in (46). Some examples are given in (47).

(46)

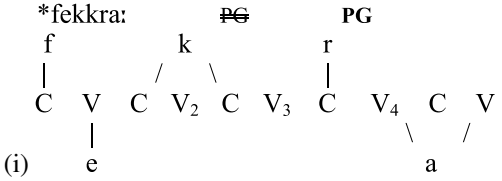
	C_0	V ∼ Ø alternation	Phonological status / V_V
(a)	{t, k, w, ʃ, tʃ}	Never	Always long
(b)	{b, d, g, ɖ, l, m, n, r, h, ʔ, ʕ, ħ}	Always	Always short
(c)	{s, f, q, j}	Sometimes	Short and long

(47)

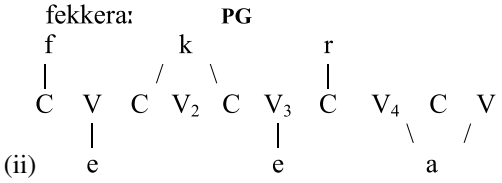
	IMP 2S	PRES 1S		Phonological status / V_V
(a)	hitiq	hitiqaː	‘walk slowly’	long
	feker	fekeraː	‘think’	
	sawaħ	sawaħaː	‘shout’	
(b)	koboʕ	kobʕaː	‘grow’	always short
	qaraq	qarqaː	‘sink’	
	faħal	faħlaː	‘plant’	

[23] In Government Phonology, this is accounted for as follows: V_3 is empty and as such cannot properly govern V_2, see (i). V_3 is identified by *e*, and all empty V positions are licensed, see (ii).

(i)



(ii)



(c)	tifiq	tifqaː	‘drip’	short
	tafaq	tafaqaː	‘sting’	long
	boqor	boqraː	‘crown’	short
	qoqob	qoqobaː	‘keep separate’	long
	qosol	qoslaː	‘laugh’	short
	fasaħ	fasaħaː	‘permit’	long
	sajaħ	sajaħaː	‘get damp’	long

It is thus clear that intervocalic [s] and [j] can be either phonologically simple or geminate: 2S *karisaː* and 1S *karijaː* exemplify /SS/ and /jj/, respectively.

To conclude, we argued that the allomorphic material in the column labeled ‘?’ in (17) above results from the addition of a CV unit by CAUS. It is fully predictable on the basis of representation (34). As such, allomorphy in class 2a does not qualify as a classificatory property.

3.3 *The representation of* -eː

3.3.1 *Class 2b verbs*

We now turn to class 2b verbs and show that class 2b and class 2a are one and the same group. According to Saeed (1993: 61ff.; 1999: 74ff.) and Orwin (1995: 60ff.), class 2b verbs are formed by adding the suffix *-eː* to nouns and adjectives:

(48) (a) Noun > class 2b

ʕaro	‘anger’	ʕareː	‘anger, make angry’
ʕaʃo	‘dinner’	ʕaʃeː	‘have dinner’
bijo	‘water’	bijeː	‘add water to s.t., dilute’

(b) Adjective > class 2b

adag	‘hard, strong’	adkeː	‘make strong, harden’
af	‘sharp point’	afeː	‘sharpen’
ʕad	‘white’	ʕaddeː	‘whiten’

Observe the data in (49).

(49)

Noun/adj	Class 2b	Class 3a
adag	adkeː	adkajso
‘hard, strong’	‘make strong, harden’	‘make strong for o.s., resist’
af	afeː	afajso
‘sharp point’	‘sharpen’	‘sharpen for o.s.’
bijo	bijeː	bijajso
‘water’	‘add water to s.t., dilute’	‘dilute for o.s.’
ʕad	ʕaddeː	ʕaddajso
‘white’	‘whiten’	‘whiten for o.s.’

Class 2b verbs are the base of the derivation of class 3a verbs: *-eː* is replaced by *-aj-* and the suffix *-so* is added. This situation is similar to that described for the derivation of class 3a verbs on the basis of class 2a verbs in (23) and (24) above.

Class 3a verbs have an additional autobenefactive value. We take this observation to indicate that class 3a verbs are built on class 2b verbs by suffixation of autobenefactive *-o*. In other words, *-ajs-* is the allomorph of class 2b *-eː* before AUTOBEN. The table in (50) shows the parallelism between class 2a and class 2b. Class 2b has *-aj-* wherever class 2a has *-i-*, with a phonologically conditioned alternation: LEX of class 2b surfaces as [eː] / __ {#, V}, and as [aj] / __{C}.

(50) *LEX in classes 2a and 2b*

		Class 2a	Class 2b
Inflection	IMP 2S	i	eː
	1S/3MS/3P (= before V)		
	2S/3FS/2P/1P (= before C)		aj
Derivation	Before autobenefactive *-o*, stative *-an*, etc.	(i)s	ajs

3.3.2 *The decomposition of* -eː

Let us start with the phonological makeup of class 2b *-eː*. According to Element Theory, *e* results from the fusion of Elements A and I: *e* = | A I|. The representation of *-eː/aj* is thus as follows:

(51) (a) CVCV (b) CVCV

```
  \ /        |  |
  A I        A  I
  [eː]       [aj]
```

We propose to analyze class 2b *-eː* as causative /It/ + floating |A|. In other words, class 2b verbs differ from class 2a verbs only with respect to the presence of |A|.

(52)

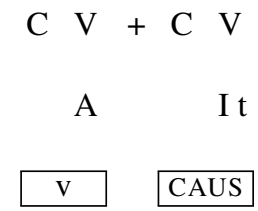
```
   C  V

A     I t
  CAUS
```

This representation immediately raises two questions: Why do class 2b verbs have an additional |A|? and Where does the extra CV position in (51) come from?

Recall that *-eː* selects for nouns and adjectives and transforms the item it attaches to into a verb. We therefore propose to analyze the additional |A| in class 2b as the exponent of verbality (v in (53)–(58)).[24] This exponent brings its own CV unit, and the final representation of *-eː/aj* is as follows:

[24] An alternative hypothesis would be to claim that |A| is a nominalizer, which attaches to adjectives to create nouns. For instance, from the adjective *ħun* 'bad' we can derive the

(53)



The item *e:/aj* differs from CAUS in the presence of the exponent of verbality only. Our analysis correctly describes the fact that

- the only way to derive a causative verb from an adjective is to add *-e:* as *-i* cannot be used in this case;
- it is impossible to derive an autobenefactive from an adjective, e.g. *adag* 'strong, hard' > **adko* (class 3b) is ungrammatical.

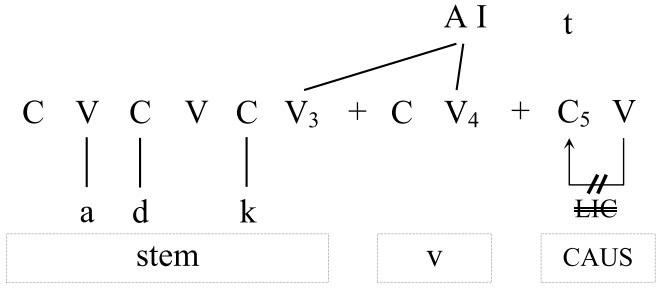
This is precisely because |A| is necessary in the derivation.

3.3.3 *Causative allomorphy (I): Derivational morphology*

The representation of a class 2b verb involves the root, the verbalizer |A| and the CAUS suffix.

Consider first the case where class 2b *-e:* is not followed by another derivational suffix in (54).

(54) |A| + *CAUS in isolation:* adke:*'make strong, harden'*



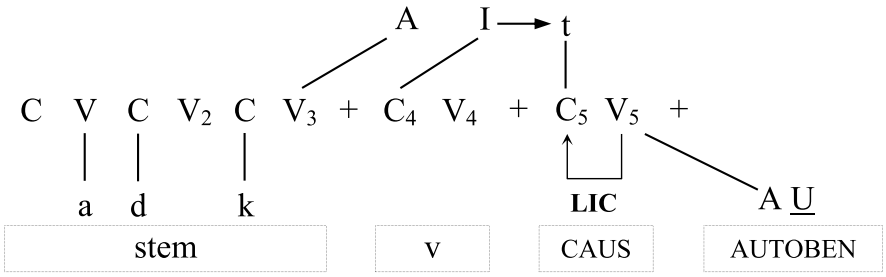
The causative *t* cannot be associated because its potential docking site, C_5, is not licensed. |A| and |I| fuse and associate with V_3 and V_4. The final empty CV, being unidentified, is not taken into account in the linearization process.

If CAUS is followed by the autobenefactive suffix, class 2b surfaces as *-ajs-* (55).

---

noun *ħuma-ha* 'the badness', and the causative verb *ħume:* 'wrong s.o.' (additional examples include *ʕad* 'white' > *ʕadda-ha* 'the white one' and *ʕadde:* 'whiten'; *san* 'good' > *sama-ha* 'the goodness' and *same:* 'prepare'). In Somali, nominalizers generally bring their own CV (Godon 1998, Barillot 2002). Since the exact grammatical status of |A| does not pertain to the representation and segmentation of *e:/aj*, we leave this debate open for further research. See Bruno (1984) for additional details.

(55) |A| + *CAUS* + *AUTOBEN:* adkajso *'make strong for s.o.'*



The suffix *-o* on the final V position (V_5) licenses *t* to be associated with C_5. As in (40) and (41) above, |I| palatalizes *t* to [s]. At the present time we have no explanation for the fact that class 2b suffix surfaces as *-aj-* when followed by C(V). We encode this observation by associating |A| to V_3 and |I| to V_4 instead of fusing |A| and |I| as in (54) above.[25] Since V_4 is properly governed by V_5, it does not need to be identified. As a result, an entire CV unit (C_4V_4) remains empty. We assume such a structure is not well-formed: C_4V_4 spells out verbality (it is the exponent of the v head), it must be identified and |I| associates with C_4.[26]

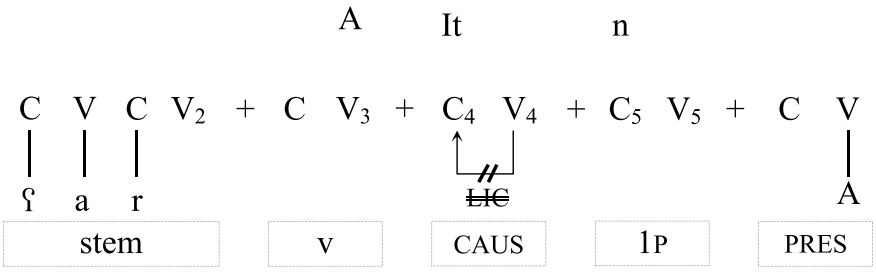
3.3.4 *Causative allomorphy (II): Inflectional morphology*

We now turn to the representations of the inflected forms and take as a representative example the verb *ʕare:* 'anger, make angry'. As above for class 2a, we focus on 1S, 2S and 1P.

We start with 1P. The basic ingredients are the root *ʕar*, the verbalizer |A|, the causative marker /It/, PNG *-n-* and TAM |A|. The underlying sequence is shown in (56a), whereas, in (56b), we illustrate the linearization and the computation of the surface form.

(56) *Class 2b, 1P, e.g.* ʕarajnna:*'we make s.o. angry'*

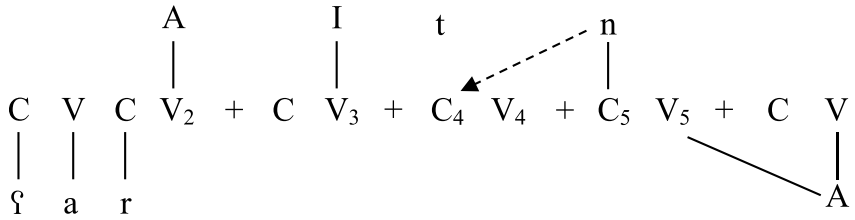
(a)



[25] The theory does not predict whether two Elements combine or not in configurations like those in (54) and (55). Two logical possibilities exist and they are both attested in Somali. At the present stage, we cannot predict why |A| and |I| combine in [adke:] 'make strong, harden' while they do not in [adkajso] 'make strong for o.s., resist'. However, we do predict that the suffix must be long, i.e. it occupies two skeletal positions.

[26] There are certainly other technical options making it possible to derive the structure in (55). The one we propose derives the correct form *adkajso* 'make strong for o.s.' and we thus leave this question for further research.

(b)

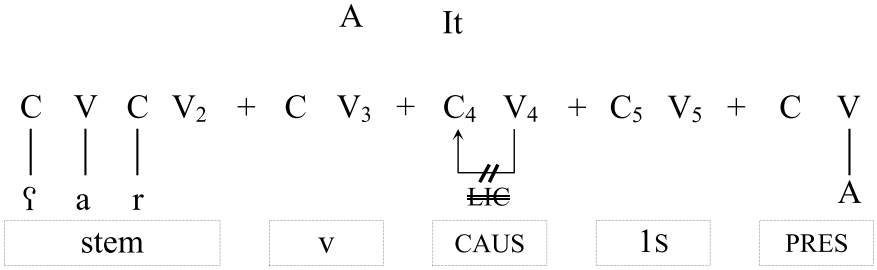


The process is parallel to that illustrated for 2a causatives in (42) above. The V position of CAUS, V_4, is empty and, as a consequence, (i) CAUS *t* is not licensed and cannot associate with C_4, and (ii) V_2 and V_3 are not properly governed. |A| associates with V_2, whereas |I| associates with V_3. *n* identifies C_4V_4 by propagation on C_4.

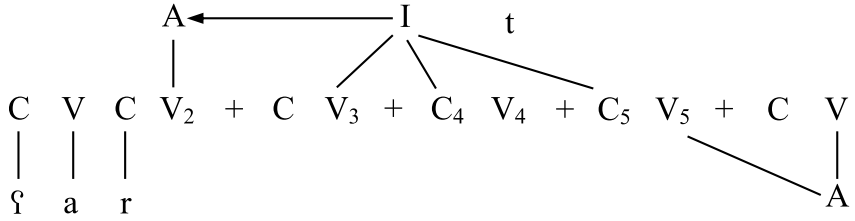
The representation of 1S *ʕare:j(j)a:* is shown in (57): the form after concatenation of the relevant morphemes appears in (57a), and (57b) shows how the segmental make up of the form is associated with the skelettal tier, yielding to phonetic interpretation.

(57) *Class 2b, 1S, e.g.* ʕare:j(j)a: *'I make s.o. angry'*

(a)



(b)

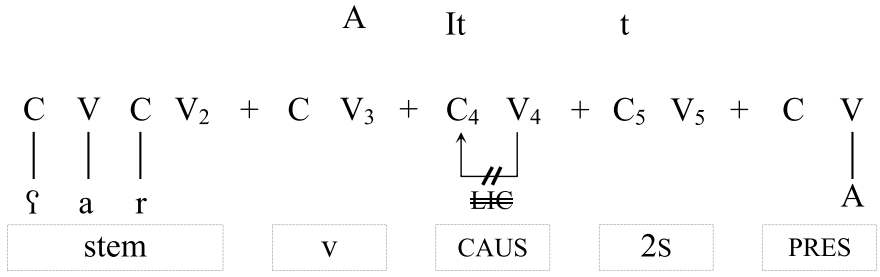


V_4 is empty. As a consequence, CAUS *t* cannot associate with C_4. V_3 and V_2 are not phonologically governed. |A| associates with V_2, whereas |I| associates with V_3. We are left with an empty $C_4V_4C_5$ sequence. The only possible repairing strategy is the propagation of |I| to C_4 and C_5. |I| palatalizes preceding |A| into surface [e].

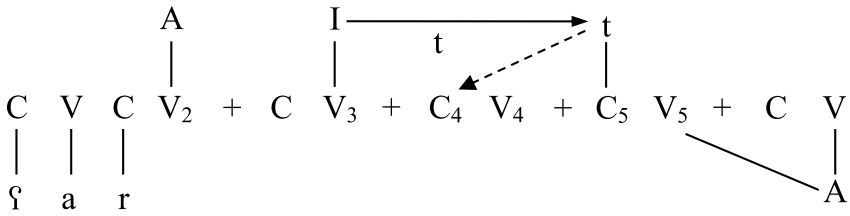
Finally, 2S is made of the same ingredients, with the exception of the person/number marker, which is *-t-*, as can be seen in (58).

(58) *Class 2b, 2S, e.g.* ʕarajsaː*'you make s.o. angry'*

(a)



(b)



In this configuration, again, V_4 is empty and does not license C_4. As a consequence, CAUS *t* cannot associate with its template. By contrast, since V_5 is identified by TAM A, 2S *t* does associate with C_5 and spreads onto C_4. Finally, |I| palatalizes underlying *t*, which surfaces as [s].

3.4 *Conclusion*

In this section, we have shown the complexity of the alternations involved in the two causative verb classes. We argued that the underlying representation of the causative marker is /It/ – recall (34) above. Regular phonological processes apply to the concatenation of the base and this suffix, which is linearized into surface forms through the principles of Government Phonology. We propose to analyze *-eː* of class 2b as involving the same causative marker /It/, plus an additional |A|, which, as we argue in Section 3.3.2, is a verbalizer. As a consequence, there is no need for a distinction between three inflectional verb classes 1, 2a and 2b.

4. CLASSES 3A AND 3B

We now turn to the analysis of classes 3a and 3b, the two 'autobenefactive' classes. Our aim is to propose a representation that accounts for the various surface realizations of the autobenefactive suffix and show that the allomorphy observed in these classes can be reduced to regular phonological processes, hence there is no need for the assumption of a specific 'autobenefactive' class.

4.1 *The autobenefactive marker*

The analysis presented in this section is based on Barillot (2002) and Barillot & Ségéral (2005). Autobenefactive verbs are divided into two classes depending on the form of the suffix attached to the basic class 1 verb, *-so* (class 3a) or *-o* (class

3b). As shown in Section 3.1 above, class 3a verbs are derived from causative verbs and *-so* must be analyzed as CAUS *-s-* + AUTOBEN *-o*. Class 3b verbs are derived from class 1 verbs by suffixation of AUTOBEN *-o*. Given this analysis, the allomorphs of AUTOBEN appear in bold face in (59).

(59)

		(a) Class 1	(b) Class 3a	(c) Class 3b
Imperative 2S		mar	mar-s-**o**	mar-**o**
Infinitive		mar-i	mar-s-**an**	mar-**an**
Progressive (present)	1S	mar-aj-aː	mar-s-**an**-aj-aː	mar-**an**-aj-aː
Present	1S	mar-aː	mar-s-**ad**-aː	mar-t-aː
	2S	mar-t-aː	mar-s-**a**-t-aː	mar-**a**-t-aː
	3MS	mar-aː	mar-s-**ad**-aː	mar-t-aː
	3FS	mar-t-aː	mar-s-**a**-t-aː	mar-**a**-t-aː
	1P	mar-n-aː	mar-s-**an**-n-aː	mar-**an**-n-aː
	2P	mar-t-aː-n	mar-s-**a**-t-aː-n	mar-**a**-t-aː-n
	3P	mar-aː-n	mar-s-**ad**-aː-n	mar-t-aː-n
		‘pass, tie up’	‘finish off, consume, dress up’	‘be finished/empty, become used up’

The suffix *-o* appears only in the imperative 2S, the citation form of verbs in Somali dictionaries and grammars. The various allomorphs of AUTOBEN, presented above in (20) and (21), are recast in (60): we abstract away from CAUS *-s-* in class 3a and filter out PNG morphemes.

(60) *AUTOBEN allomorphy*

		Class 3a	Class 3b
Inflection	Imperative 2S	o	o
	1S/3MS/3P (= before V)	ad	t
	2S/3FS/2P (= before C except *n*)	a	a
	Infinitive/progressive/1P (= before *n*)	an	an
Derivation	Before stative *-an*, etc.	Does not apply	

This leads us to include both *-d-* (1S/3MS/3P, class 3a) and *-n-* (infinitive/progressive/1P, both classes) into LEX. AUTOBEN is composed of a vowel *-a-*, which alternates with Ø in class 3b (1S/3MS/3P), and of a coronal consonant (*t*, *d* or *n*). This exhausts the possibilities since no other derivational suffix may appear to the right of AUTOBEN.

The table in (60) reveals that AUTOBEN displays a very similar allomorphy in both classes: the only difference lies in 1S/3MS/3P forms, which are shaded in gray. There, two differences are found: (i) the vowel *-a-* appears only in class 3a, and (ii) class 3a has *-d-* while class 3b has *-t-*.

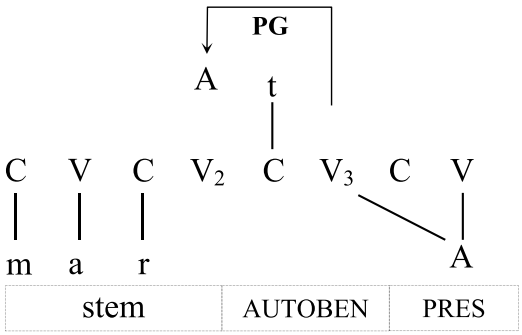
Let us start with the a ∼ Ø alternation, e.g. 1S class 3a *marsa̲daː* ‘I finish off, consume, dress up’ ∼ class 3b *martaː* ‘I am finished/empty, become used up’. This alternation has to be compared with that in (61), involved in CVCVC class 1 verbs, e.g. *gudub* ‘cross’ (recall example (10a) above):

(61)	(a)	$CV_iCV_iC\#$	gudub	/gudb/	IMP 2S
	(b)	$CV_iCCV\ldots$	gudbaː	/gudb-Ø-aː/	PRES 1S/3MS
			gudbaːn	/gudb-Ø-aː-n/	PRES 3P
	(c)	CV_iCV_iCCV	gudubtaː	/gudb-t-aː/	PRES 2S/3FS
			gudubnaː	/gudb-n-aː/	PRES 1P
			gudubtaːn	/gudb-t-aː-n/	PRES 2P

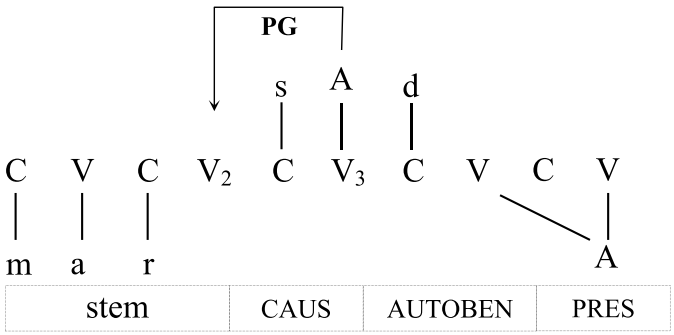
This alternation stems from the fact that CCC and CC# clusters are prohibited in Somali. To avoid such clusters, the preceding stem vowel *u* propagates in the relevant contexts. The a ∼ Ø alternation in AUTOBEN has the same cause: *-a-* surfaces to avoid either a *CCC or a *CC# cluster. In class 3b 1S, *mar-t-aː* 'I am finished/empty, become used up' does not contain any putative CCC cluster and *-a-* does not appear (**mar-at-aː*). In class 3a 1S, by contrast, *mar-s-ad-aː* 'I finish off, consume, dress up' contains a putative CCC cluster, **marsd-aː*, and *-a-* shows up. In the framework of Government Phonology, this type of V ∼ Ø alternation regularly derives from Proper Government relations: *a* does not surface in *martaː* because V_2 is properly governed by V_3, as seen in (62a).

(62) *Class 3a/b, 1S*

(a) *Class 3b, 1S, e.g.* martaː *'I am finished/empty'*



(b) *Class 3a, 1S, e.g.* marsadaː *'I finish off, consume, dress up'*



By contrast, in /mar-s-d-aː/ in (62b), either V_2 or V_3 has to be identified in order for the representation to be well-formed. Of these two conceivable options, it is the second one which is applied.[27]

We conclude that the a ∼ Ø alternation opposing class 3a/b verbs is strictly conditioned by the phonological context: no unpredictable allomorphy is involved.[28]

We can now turn to the representation of AUTOBEN. The alternating vowel in class 3a/b is *a*. We propose that this | A| belongs to AUTOBEN. Moreover, since it alternates with Ø, it has to be considered as a floating vowel. NOW let us turn to the identity of the underlying consonant: it must be either *d* or *t*. We propose it is *t*, and the *t* ∼ *d* alternation derives from a general voicing rule stated in (9) above, /t/ → [d] / V_V:

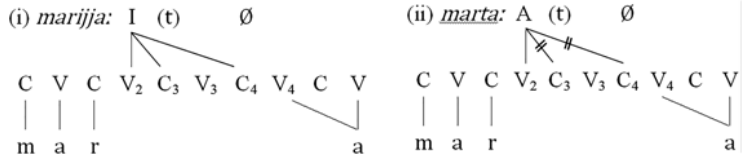
(63) AUTOBEN CV
A t

To sum up: the alleged allomorphy in class 3a/b 1S/3MS/3P forms (*marsadaː* 'I finish off, consume, dress up' vs. *martaː* 'I am finished/empty, become used up') is accounted for by general phonological rules. The underlying representation of AUTOBEN in (63) is the same in both classes. It includes two parts: *t*, which is regularly voiced in intervocalic position, and | A|, which regularly surfaces when its absence would lead to a prohibited consonant cluster.

At first sight, this proposal seems to be immediately contravened by the 2S/3FS/2P forms of both classes. In these forms, (i) -*t*- surfaces in intervocalic position, and (ii) in class 3b, the floating vowel -*a*- surfaces even if it is preceded by a single consonant: its appearance cannot be ascribed to the necessity of preventing a prohibited CCC cluster:

---

[27] A comparison of *martaː* (1S class 3b) and *marijjaː* (1S class 2a) reveals that *martaː* is shorter than *marijjaː*. This is due to an asymmetry between |I| and |A|: while |I| may propagate onto C_3 and C_4 and surface as [jj] in *marijjaː* as in (i), | A| may not propagate to C positions, as in (ii).



[28] This analysis receives strong support from another fact. The very same alternation obtains WITHIN class 3b. If AUTOBEN -*o* is added to a CVCVC class 1 verb, the second vowel does not appear: *barah* 'dilute s.t. with water' > *barho* 'mix milk with water for o.s.'. The 1S present of *barho*, /barh-t-aː/ (or /barh-d-aː/), contains a putative CCC cluster. We expect -*a*- to surface, and this is the case: the attested form is *barhadaː*. For more details, see Barillot & Ségéral (2005).

(64) Present 2S

Class 1	Class 3a	Class 3b
mar-t-a: 'you tie up/pass'	mar-s-a-t-a: 'you finish off, consume, dress up'	mar-a-t-a: 'you are finished/empty, become used up'
	Intervocalic [t]	
	No putative CCC	

However, as we turn to show, the second *-a-* in *marata:* does appear in the relevant context: C_CC.

4.2 *Virtual geminates*

The underlying structure of *marata:* 'you tie up for yourself' is as follows:

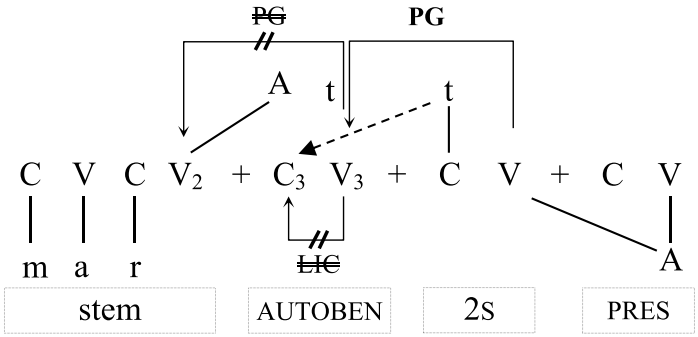
(65)

mar	at	t	a:
'tie up'	AUTOBEN	2S	TAM

The presence of *-a-* is accounted for: it is due to the /tt/ cluster to its right. The floating vowel *-a-* surfaces to break up a CCC cluster (r-t-t). Moreover, it explains why intervocalic [t] is not voiced: it is a geminate. In other words, /t/ → [d] / V_V, and /tt/ → [t] / V_V.[29] The representations of class 3b, 2S *marata:* and class 3a, 2S *marsata:* are given in (66a) and (66b), respectively.

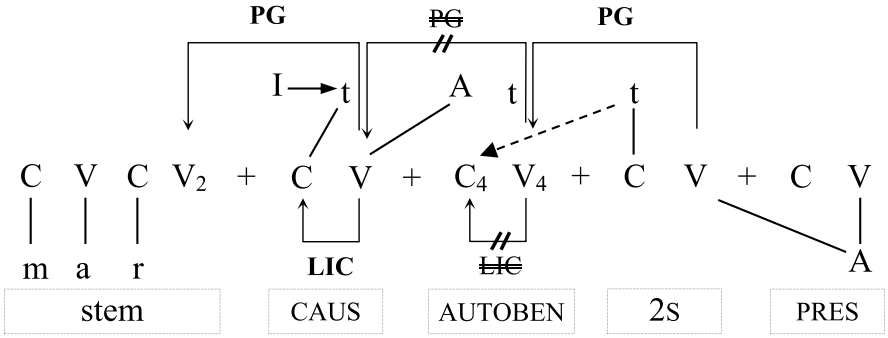
(66) *Class 3a/b, 2S*

(a) *Class 3b, 2S, e.g.* marata: *'you are finished/empty, become used up'*



[29] The analysis of intervocalic [t] as /tt/ makes it possible to explain several facts of Somali morphology (see Barillot 2002, Barillot & Ségéral 2005). For instance, in CVtVC verbs, the second vowel is always present (as seen in example (47) above).

(b) *Class 3a, 2S, e.g.* marsataː *'you finish off, consume, dress up'*



An interesting by-product of this analysis comes from the examination of a particular group of class 3b verbs. Recall from (46) above that intervocalic *k*, *ʃ* and *w* are underlying geminates, like *t*. We predict that if CV*ko*, CV*ʃo*, CV*wo* autobenefactive verbs exist, they will not behave like *maro* in (67a), but rather like *marso* in (67b) because their last consonant is phonologically long. This is exactly what we observe in (67c):

(67)

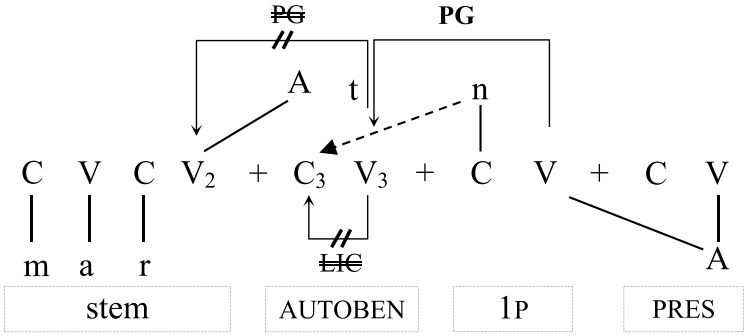
	IMP 2S	PRES 1S No -a-	-a-	
(a)	maro	marta:		'be finished/empty, become used up'
(b)	marso		marsada:	'finish off, consume, dress up'
(c)	tuko		tukada:	'pray'
	buko		bukada:	'become sick'
	ɖako		ɖakada:	'hide o.s.'
	ʕeʃo		ʕeʃada:	'keep for o.s.'
	fuʃo		fuʃada:	'achieve for o.s.'
	gaʃo		gaʃada:	'wear'
	duwo		duwada:	'bypass for o.s.'
	huwo		huwada:	'cover o.s. with s.t.'
	daːwo		daːwada:	'watch'

4.3 *1P and imperative 2S*

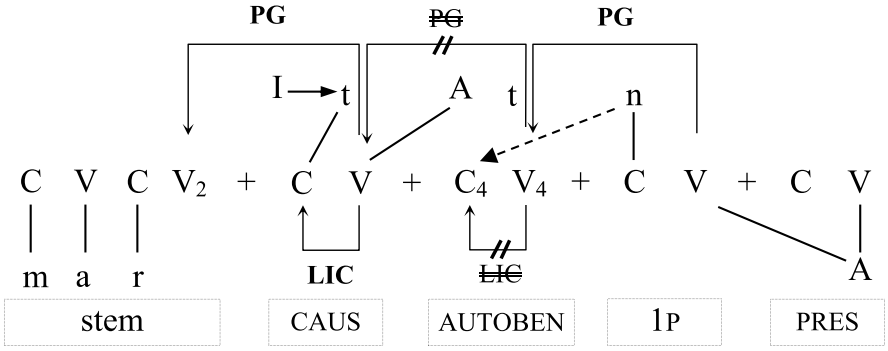
We are left with two cases of allomorphy: (i) 1P in which AUTOBEN surfaces as *-an-*, and (ii) 2S IMP in which AUTOBEN surfaces as *-o*. Let us first consider *-an-* in 1P (*maranna:* 'are finished/empty, become used up'), infinitive (*maran* 'be finished/empty, become used up') and progressive forms (*maranaja:* 'I am becoming used up'). The underlying structure of 1P can be broken down into four parts, and represented as in (68a): V_3 being empty, it does not properly govern V_2 nor does it license C_3.

(68) *Class 3a/b, 1P*

(a) *Class 3b, 1P, e.g.* maranna: *'we are finished/empty, become used up'*



(b) *Class 3a, 1P, e.g.* marsanna: *'we finish off, consume, dress up'*



This has two consequences: (i) V_2 must be identified – AUTOBEN |A| associates with this position, and (ii) *t* cannot be associated with C_3. 1P *-n-* spreads and identifies C_3. (68b) illustrates the same 1P form for class 3a *marsanna:*. The only difference lies in the presence of CAUS to the left of AUTOBEN, which surfaces as [s].

Consider now the 2S imperative. Why does AUTOBEN surface as *-o* in this form, and only in this form? In order to answer this question, we have to understand the distribution of final *-o*s in Somali. For that purpose, consider the plural marker *-o* in (69a–c).

(69)

	SG	PL	SG + DET	PL + DET	
(a)	na:g	na:go		na:ga-ha	'woman'
(b)	inan	inammo		inamma-da	'boy'
(c)	ilig	ilko		ilka-ha	'tooth'
(d)	ho:jo		ho:ja-da		'mother'

When the determiner (either [ha] MASC, or [da] F) is suffixed to a plural noun ending in *-o*, this vowel alternates with *-a-*, as seen in (69a–c). The same alternation takes place in singular nouns ending in *-o*, seen in (69d).

We take this alternation to indicate that final short *-a* is banned in Somali:

(70) *a# in Somali

This is confirmed by an exhaustive examination of Agostini et al. (1985): nouns with a final short vowel overwhelmingly end in e#, i# or o#, as anticipated by (70).[30]

(71)

a#	e#	i#	o#	u#
27	991	1015	2390	2

(Agostini et al. 1985)

Returning to class 3b, we observe that the 2S imperative is the only form where AUTOBEN appears in word-final position. In all other cases, AUTOBEN is followed by either PNG or TAM. We thus claim that *-o* surfaces in 2S imperative to satisfy (70). In Element Theory, [o] and [a] share Element |A|, and differ in that [o] contains |U|:

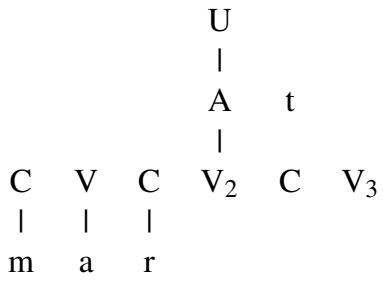
(72) (a) [a] = |A|

(b) [o] = |U A|

In class 3 verbs, exactly as in nouns, an additional |U| appears in order to satisfy (70): this is again a general phonological rule of Somali, and not an idiosyncratic property of class 3a/b.

The representation of the imperative 2S *maro* (class 3b) appears in (73).

(73) *Imperative, 2S (class 3b):* maro *'be finished/empty!'*



We predict that *-t* does not surface: V_3 being empty, it cannot license the preceding C position, which would be the only C-slot available to *-t*.

4.4 *Conclusion*

In this section, we have pursued our analytical path into the complex patterns of classes 3a and 3b. We observed that, abstracting away from *-s-* in class 3a,

[30] As pointed out by an anonymous *JL* referee, Somali has a few examples of words with final *-a*, e.g. *laba* 'two', *toddoba* 'seven', etc. In these cases, however, final *-a* may also be realized as *-o*. Next to these cases, there are two grammatically defined contexts in which final *-a* surfaces: the reduced paradigm verb endings (e.g. *ana: tagayá* 'I am going') and the definite article *-ka* (M) and *-ta* (F). They merit being listed and analyzed specifically. In particular, note that final *-á* of the reduced paradigm is clearly a reduction of final long *-a:* (see Saeed 1993: 106, who says: 'In present tense forms ending in *-a:* . . . the ending *-a:* is . . . reduced to *-á*')'. We leave this question open for further research.

which is CAUS, class 3a and class 3b behave on a par, except for one property: in the 1S/3MS/3P, class 3a displays LEX = *ad*, whereas class 3b has LEX = *t*. We proposed an underlying representation for AUTOBEN and showed that the diverging patterns in class 3a and 3b can be derived from this representation by regular phonological principles. Finally, we showed that the inflected forms of both classes can be represented on a par with the inflected forms of class 1 and class 2.

5. CONCLUSION AND FURTHER ISSUES

We have argued that the allomorphic patterns targeting the complex derivational and inflectional morphological processes of Somali verbs are epiphenomenal. More precisely, we reduced the allomorphy exhibited by LEX and PNG markers to a case of phonologically conditioned allomorphy. Our approach has a straightforward impact on the organization of Somali verb classes: there is no need to divide Somali verbs into 3 distinct conjugations. We claim that surface distinctions are made exclusively on the basis of the exponent selected by each verbal root.

We have established unified representations of the causative and the autobenefactive markers (see (34) and (63) above). These representations are repeated below. Our central claim is that all surface realizations of CAUS and AUTOBEN derive from these representations.

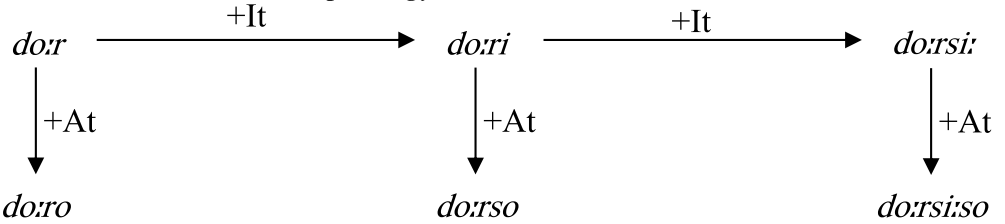
(74)

C V	C V
I t	A t
CAUS	AUTOBEN

On the one hand, the representations in (74) share some phonological material, namely the CV unit and the consonant /t/. On the other hand, they differ with respect to the vocalic Element: caus contains |I|, whereas AUTOBEN contains |A|. We propose to interpret the phonological material shared by both CAUS and AUTOBEN as the exponent of a derivational marker. This marker is responsible for building a new word from a well-formed noun, verb or adjective, either of the same category or of a different one. As for the Elements, they are the actual causative and autobenefactive exponents.

The possible paths of verbal formation are schematized in (75).

(75) *Somali derivational morphology*



Besides the combination of CAUS and AUTOBEN examined in this article, our analysis accounts for two additional combinations of suffixes: CAUS + CAUS

('strong causative' in *-siː*, Saeed 1993: 59–60) and CAUS + CAUS + AUTOBEN (*-siːso*).

Finally, in our approach, we dispense with the notion of paradigm as a linguistic active object. Paradigms result as the output of phonological computation and do not need to be specified in advance in the lexical entry of each verb. Further research, drawing on additional Somali and cross-linguistic data will have to corroborate this hypothesis on the status of paradigms.

REFERENCES

Agostini, Francesco, Annarita Puglielli & Ciise Mohamed Siyaad (eds.). 1985. *Dizionario Somalo Italiano [Somali–Italian dictionary]*. Roma: MAE, Dipartimento per la Cooperazione allo Sviluppo.

Anderson, Stephen. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Andrzejewski, BogumiłWitalis. 1964. *The declensions of Somali nouns*. London: School of Oriental and African Studies.

Andrzejewski, BogumiłWitalis. 1969. Some observations on hybrid verbs in Somali. *African Language Studies* 10, 47–89.

Andrzejewski, BogumiłWitalis. 1979. *The case system in Somali*. London: School of Oriental and African Studies.

Aronoff, Mark. 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.

Aronoff, Mark. 1994. *Morphology by itself*. Cambridge, MA: MIT Press.

Backley, Philip. 2011. *An introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Banti, Giorgio. 1988. Two Cushitic systems: Somali and Oromo. In Harry van der Hulst & Norval Smith (eds.), *Autosegmental studies on pitch accent*, 11–49. Dordrecht: Foris.

Banti, Giorgio. 2012. Les parties du discours en somali. Presented at Xuska 40 guurada farsoomaliida [Celebrations for the 40th anniversary of Somali writing], Université de Djibouti.

Barillot, Xavier. 2002. *Morphophonologie gabaritique et information consonantique latente en somali et dans les langues est-couchitiques*. Ph.D. dissertation, Univerité Paris 7.

Barillot, Xavier & Sabrina Bendjaballah. 1998. Some aspects of verbal derivational morphology in Somali: Remarks on the 'causative' conjugation. Presented at the 28th Colloquium on African Languages and Linguistics, Leiden, 31/08-2/09/1998.

Barillot, Xavier & Philippe Ségéral. 2005. On phonological processes in the '3rd' conjugation in Somali. *Folia orientalia* 41, 115–131.

Bendjaballah, Sabrina. 1998. La palatalisation en somali. *Linguistique Africaine* 21, 5–52.

Bendjaballah, Sabrina. 1999. *Trois figures de la structure interne des gabarits*. Ph.D. dissertation, Univerité Paris 7.

Bendjaballah, Sabrina & Martin Haiden. 2008. A typology of emptiness in templates. In Hartmann et al. (eds.), 21–57.

Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 42, 531–573.

Blevins, James P. 2015. Inflectional paradigms. In Matthew Baerman (ed.), *The Oxford handbook of inflection*, 87–111. Oxford: Oxford University Press.

Bonet, Eulàlia & Daniel Harbour. 2012. Contextual allomorphy. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, 195–235. Oxford: Oxford University Press.

Bruno, Biancamaria. 1984. Note sui verbi di derivazione nominale in somalo [Remarks on denominal verbs in Somali]. In Puglielli (ed.), 113–131.

Carstairs-McCarthy, Andrew. 1998. Paradigmatic structure: Inflectional paradigms and morphological classes. In Andrew Spencer & Arnold Zwicky (eds.), *The handbook of morphology*, 322–334. Oxford: Oxford University Press.

Carstairs-McCarthy, Andrew. 2005. Affixes, stems and allomorphic conditioning in paradigm function morphology. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology 2005*, 253–281. Dordrecht: Springer.

Charette, Monik. 1989. The Minimality Condition in phonology. *Journal of Linguistics* 23, 213–243.

Charette, Monik. 1990. Licence to govern. *Phonology* 7.2, 233–253.

Charette, Monik. 1991. *Conditions on phonological government*. Cambridge: Cambridge University Press.

Cristófaro Silva, Thaïs. 2003. Palatalisation in Brazilian Portuguese. In Stefan Ploch (ed.), *Living on the edge: 28 papers in honour of Jonathan Kaye*, 243–257. Berlin: Mouton de Gruyter.

Embick, David. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.

Godon, Elsa. 1998. Aspects de la morphologie nominale du somali: la formation du pluriel. MA dissertation, Université Paris 7.

Goldsmith, John. 1978. *Autosegmental Phonology*. New York: Garland Press.

Guerssel, Mohand. 1990. On the syllabification pattern of Berber. Ms, Université du Québec à Montréal.

Guerssel, Mohand. 2003. The metathesis effect in Classical Arabic and the representation of geminates. In Jacqueline Lecarme (ed.), *Research in Afroasiatic grammar 2* (Current Issues in Linguistic Theory 241). Amsterdam: John Benjamins.

Hartmann, Jutta, Veronika Hagedus & Henk van Riemsdjik (eds.). 2008. *The sound of silence: Empty elements in syntax and phonology*. Amsterdam: Elsevier.

Kang, Yoonjung. 2011. Loanword phonology. In Marc van Oostendorp, Ewen Colin, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2258–2281. New York: Wiley-Blackwell.

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*. Firenze: E. Ariani.

Kiparsky, Paul. 1996. Allomorphy or morphophonology? In Rajendra Singh (ed.), *Trubetzkoy's orphan. Proceedings of the Montréal Roundtable Morphophonology: Contemporary Responses*, 12–31. Amsterdam & Philadelphia, PA: John Benjamins.

Leben, Williams. 1973. *Suprasegmental phonology*. Ph.D. dissertation, MIT.

Lowenstamm, Jean. 1996. CV as the only syllable type. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology*, 419–441. Manchester: European Studies Research Institute, University of Salford.

Lowenstamm, Jean. 2008. On little n, $\sqrt{}$ and types of nouns. In Hartmann et al. (eds.), 105–143.

Orwin, Martin. 1995. *Somali*. New York: Routledge.

Puglielli, Annarita (ed.). 1984. *Aspetti morfologici, lessicali e della focalizzazione* [Differents perspectives on the morphology, the lexicon and focus] (Studi Somali 5). Roma: MAE, Dipartimento per la Cooperazione allo Sviluppo.

Puglielli, Annarita & Cabdallah C. Mansuur. 2012. *Qaamuuska Af-Soomaaliga*. Roma: UniTrePress.

Puglielli, Annarita & Ciise M. Siyaad. 1984. La flessione del nome [Noun inflection]. In Puglielli (ed.), 53–112.

Saeed, John Ibrahim. 1993. *Somali reference grammar*, 2nd revised edn. Kensington, MD: Dunwoody Press.

Saeed, John Ibrahim. 1995. The semantics of middle voice in Somali. *African Languages and Cultures* 8.1, 61–85.

Saeed, John Ibrahim. 1999. *Somali*. Amsterdam & Philadelphia, PA: John Benjamins.

Scheer, Tobias. 2004. *A lateral theory of phonology,* vol. 1: *What is CVCV, and why should it be?* Berlin: Mouton de Gruyter.

Scheer, Tobias & Philippe Ségéral. 2001a. La coda miroir. *Bulletin de la Société Linguistique de Paris* 96, 107–152.

Scheer, Tobias & Philippe Ségéral. 2001b. Fake palatalizations. Presented at the 33rd Poznań Linguistic Meeting.

Ségéral, Philippe & Tobias Scheer. 2001. Abstractness in phonology: The case of virtual geminates. In Katarzyna Dziubalska-Kołaczyk (ed.), *Constraints and preferences*, 311–337. Berlin: Mouton de Gruyter.

Trommer, Jochen (ed.). 2012. *The morphology and phonology of exponence* (Oxford Studies in Theoretical Linguistics 41). Oxford: Oxford University Press.

Zorc, David R. 1993. *Somali–English dictionary with English index*. Kensington, MD: Dunwoody Press.

983 *Authors' addresses:* *(Barillot)*
Université Côte d'Azur, Laboratoire Bases, Corpus, Langage – UMR 7320-BCL, 24, avenue des Diables Bleus, 06357 Nice Cedex 4, France
xavier.barillot@unice.fr

(Bendjaballah)
Université de Nantes, Laboratoire de Linguistique de Nantes – UMR 6310-LLING, Chemin de la Censive du Tertre, BP 81227, 44312 Nantes Cedex 3, France
sabrina.bendjaballah@univ-nantes.fr

(Lampitelli)
Université François Rabelais de Tours, Laboratoire Ligérien de Linguistique – UMR 7270-LLL, 3, rue des Tanneurs, BP 410, 37041 Tours Cedex 1, France
nicola.lampitelli@univ-tours.fr
984