Somali and the Nature of Morphophonological Alternations

Daniel Bruhn
Linguistics 211B

1 Introduction

Somali is an East Cushitic language spoken in the Horn of Africa, with determiner morphemes whose initial consonants undergo interesting phonological alternations when suffixed onto noun stems. This could be described as a type of "derived environment contact phenomena," as the changes in these suffix-initial consonants are totally dependent on the final segment of the stem and are not active in other types of derived environments or within roots. In this paper I examine the alternations these determiner morphemes undergo when suffixed in different positions, sketching OT approaches to describe the phonological processes and developing a cophonological model that isolates each set of phonological alternations to its respective morphological environment.

§2 lists the segment inventory of Somali and describes the transcription system utilized in this paper. §3 presents the first set of determiner-suffix alternations, while §4 discusses some properties of morphophonological alternations that could be problematic for an OT approach. §5 examines the first set of determiner-suffix alternations, with a "brute-force" OT approach in §5.1 and a scales model in §5.2. §6 introduces the second set of alternations, which are then analyzed with a scales approach in §7. Finally, §8 sketches a method for unifying the analyses of each set of alternations into a cophonological model, and §9 concludes the paper.

2 Segment Inventory & Transcription

The Somali data in this paper is taken from Saeed (1999), Zorc & Osman (1993), and Somalism.com. All of these resources utilize the standard orthography, which does not represent tone (of which Somali has H, L, & HL). A process of [±ATR] vowel harmony exists in Somali words (and phrases, to some extent), which is also totally obscured by the orthography. In transcribing the Somali data into IPA, therefore, I have omitted tone and chosen to represent vowels and diphthongs by a minimally-modified version of the standard orthography, shown below with the segment inventory of Somali. Additional phonetic details, such as the lenition of /h/ to [ɦ] between vowels, were provided by Armstrong (1964).

Vowel & diphthong inventory:

Standard Orthography	IPA [+ATR], [–ATR]	IPA in this paper
i	i, ɪ	i
e	e, ɛ	e
a	æ, ɑ	a

o	ö, ɔ	o
u	ʉ, u	u
ay	æi, ɑɪ	aj
aw	æʉ, ɑu	aw
ey	ei, ɛɪ	ej
oy	öi (no [–ATR] version)	oj
ow	öʉ, ɔu	ow

Long vowels and diphthongs: ii ee aa oo uu aaj aaw eej ooj oow

Consonant inventory:

plosives	fricatives	other
b	f	tʃ
t (dental)	s	m
d (dental)	ʃ	n
ɖ	χ (voiceless uvular)	r
k	ħ (voiceless phary.)	l
g	ʕ (voiced phary.)	w
ɢ (voiced uvular)	h [ɦ intervocalically]	j
ʔ		

Geminate consonants:[1] bb dd ɖɖ gg ll mm nn rr

- no geminate fricatives or affricates
- /k/ and /t/ are restricted to onsets (except for a few loanwords), so no /kk/ or /tt/

3 Data: Alternations (Set #1)

The alternations in question involve a set of determiner suffixes that begin with /k/ (masc.) or /t/ (fem.), depending the gender of the head noun to which they attach:

Definite articles	Demonstratives	
-ka/ta (non-remote)	-kan/tan	‘this, these’ (close to speaker)
-kii/tii (remote)	-kaa(s)/taa(s)	‘that, those’ (further away)
	-keer/teer	‘that, those’ (middle distance)
	-koo/too	‘that, those’ (far distance)

[1] Saeed (1999) lists /qq/ as a legitimate geminate consonant, but an online dictionary search (www.somalism.com/dictionary/) yielded no words with /qq/, so I will consider it an unattested segment.

Possessives		Interrogatives
-kaj/taj	1sg ‘my’	-kee/tee ‘which?’
-kaa/taa	2sg ‘your’	
-kiis/tiis	3sg masc ‘his’	
-keed/teed	3sg fem ‘her’	
-kaja/taja	1pl exclusive ‘our’	
-keen/teen	1pl inclusive ‘our’	
-kiin/tiin	2pl ‘your’	
-kood/tood	3pl ‘their’	

When one of these determiners is suffixed onto a head noun, the suffix-initial consonant undergoes an alternation based on the final segment of the noun, surfacing as [k], [g], [h], or Ø for the /k/-initial suffixes and [t], [d], or [ɖ] for the /d/-initial suffixes. In addition, when the noun ends in /l/, a suffix-initial /t/ coalesces with it to surface as [ʃ]. These sandhi effects are detailed and exemplified below:

Surface forms of suffix-initial /k/

[k] after stem-final [b], [d], [ɖ], [f], [s], [ʃ], [tʃ], [l], [n], [r], [l]

e.g. /dab/ ‘fire’ (masc.) + /-ka/ (non-remote def.) → [dab**k**a] ‘the fire’
/gees/ ‘horn’ (masc.) + /-kan/ (near demons.) → [gees**k**an] ‘this horn’
/nin/ ‘man’ (masc.) + /-kee/ (interrog.) → [nin**k**ee] ‘which man?’

[g] after stem-final [g], [w], [j], and all vowels/diphthongs except [e], [o], [a]

e.g. /tuug/ ‘thief’ (masc.) + /-ka/ (non-remote def.) → [tuug**g**a] ‘the thief’
/ardaj/ ‘student’ (masc.) + /-kan/ (near demons.) → [ardaj**g**an] ‘this student’
/guri/ ‘house’ (masc.) + /-kee/ (interrog.) → [guri**g**ee] ‘which house?’

[ɦ] after stem-final [e], [o], [a] (with stem-vowel assimilation to suffix vowel across [ɦ])

e.g. /aabbe/ ‘father’ (masc.) + /-kaj/ (1sg poss.) → [aaba**ɦ**aj] ‘my father’
/bare/ ‘teacher’ (masc.) + /-kii/ (remote def.) → [bari**ɦ**ii] ‘the teacher (remote)’
/ilko/ ‘teeth’ (masc.) + /-ka/ (non-remote def.) → [ilka**ɦ**a] ‘the teeth’

Ø after stem-final [ɢ], [χ], [ʕ], [ħ], [ʔ], [h]

e.g. /duɢ/ ‘elder’ (masc.) + /-ka/ (non-remote def.) → [duɢa] ‘the elder’
/saʕ/ ‘cow’ (masc.) + /-ka/ (non-remote def.) → [saʕa] ‘the cow’
/tʃaah/ ‘face’ (masc.) + /-kaa/ (2sg poss.) → [tʃaaɦaa] ‘your face’

Surface forms of suffix-initial /t/

[t] after stem-final [b], [g], [f], [s], [ʃ], [tʃ], [n], [r]

e.g. /ħeeb/ 'shore' (fem.) + /-ta/ (non-remote def.) → [ħeeb**t**a] 'the shore'
/naag/ 'woman' (fem.) + /-too/ (far demons.) → [naag**t**oo] 'yon woman'
/sun/ 'poison' (fem.) + /-tan/ (near demons.) → [sun**t**an] 'this poison'

[d] after stem-final [d], [ɢ], [ʔ], [χ], [ʕ], [ɦ], [h], [w], [j], and all vowels/diphthongs

e.g. /bad/ 'sea' (fem.) + /-tee/ (interrog.) → [bad**d**ee] 'which sea?'
/bah/ 'noble' (fem.) + /-ta/ (non-remote def.) → [bah**d**a] 'the noble'
/hooyo/ 'mother' (fem.) + /-taa/ (2sg poss.) → [hooya**d**aa] 'your mother'
(w/ stem-final [o] assimilation)

[ɖ] after stem-final [ɖ]

e.g. /gabaɖ/ 'girl' (fem.) + /-ta/ (non-remote def.) → [gabaɖ**ɖ**a] 'the girl'
/feeɖ/ 'rib' (fem.) + /-ta/ (non-remote def.) → [feeɖ**ɖ**a] 'the rib'

coalesces with stem-final /l/ to form **[ʃ]**

e.g. /meel/ 'place' (fem.) + /-ta/ (non-remote def.) → [mee**ʃ**a] 'the place'
/ul/ 'stick' (fem.) + /-tee/ (interrog.) → [u**ʃ**ee] 'which stick?'

4 Discussion: Issues in Modeling Morphophonology

Having given a cursory description to these consonant changes, some might be content to treat the various surface forms of suffix-initial /k/ and /t/ as a list of allomorphs that needs no further analysis. However, given the fact that these are (mostly) single segment alternations that involve no suppletion, I believe it behooves the field of phonology to treat this phenomenon as an instance of morphologically-conditioned phonology (as opposed to phonologically-conditioned allomorphy) and to attempt a phonological explanation. Though difficult to pinpoint, there do exist phonological generalizations in the alternations (such as intervocalic voicing of /k/ and /t/) that warrant a theoretical formulation.

It should be noted, however, that this is an area in which one should be aware of a model's limitations before blindly applying it to the phenomena at hand. The phonological processes evident in the morphophonological environments here exhibit a significant difference from those that led to the development of Optimality Theory. One of the driving factors in the rise of OT was the observation of a "functional unity" among various phonological rules (Kisseberth 1970, cited in Kager 1999): several rules were observed to target the same conditioning environment, though they produce different outputs. Kager illustrates this with the following abstract example (1999, p. 56):

A set of functionally coherent rules

a. A → B / X_Y
b. A → C / X_Y
c. A → Ø / X_Y
d. Y → Z / XA_
e. Ø → B / XA_Y
f. X → Ø / _AY

The generalization uniting the above set of rules is that they are all "conspiring" to eliminate the bad sequence XAY. Different output structures result, but the functional unity of this *conspiracy* lies in the common conditioning environment XAY targeted by each rule.

OT captures such conspiracies with markedness constraints, such as *XAY, that tag specific sequences as "marked" or "in need of repair." Depending on the order of the faithfulness constraints above which *XAY is ranked, various repair strategies might be undertaken to fix the marked XAY sequence, illustrated by the range of outputs produced by rules a-f in the above example. It might be the case that one language utilizes a certain repair strategy to fix XAY, while a separate language employs a different repair strategy, or that a single language utilizes multiple repair strategies to fix XAY in different domains. The fundamental point is that a conspiracy in OT involves multiple phonological processes targeting a single conditioning environment for repair.

A type of functional unity is also evident in the morphophonological processes of the Somali data, but it is not to be found in the conditioning environment for each process. Rather, this unity exists in the use of a single repair strategy to target multiple marked sequences. A grouping of the stem-final consonants that cause suffix-initial /t/ to surface as [d] illustrates this observation:

Suffix-initial /t/ surfaces as [d] after stem-final	Possible marked environment
[d]	BP (voiced plosive followed by voiceless with the same POA)
[w], [j], and all vowels/diphthongs	VC̥V (C̥ = voiceless consonant, V includes glides & diphthongs)
[ɢ], [ʔ], [χ], [ʕ], [ħ], [h]	$C_{\text{post-velar}}$t

As shown above, the same morphophonological process of /t/ → [d] is used to repair three different marked sequences. Less dramatically, suffix-initial /k/ → [g] targets two different environments:

Suffix-initial /k/ surfaces as [g] after stem-final	Possible marked environment
[g]	BP
[w], [j], and all vowels/diphthongs except [e], [o], [a]	VC̥V

Given that a conspiracy in OT involves the association of a single markedness constraint with multiple repair strategies, the association here of a single repair strategy with multiple markedness constraints could aptly be called a "reverse conspiracy."

Since classic OT is equipped and designed to handle regular conspiracies, its applicability to reverse conspiracies is dubious. Granted, one might argue that the symmetry should be obvious: a conspiracy calls for the elevation of a single markedness constraint over a set of diverse of faithfulness constraints, while a reverse conspiracy involves the subjugation of a single faithfulness constraint below a set of diverse markedness constraints. The danger in this approach, however, comes from a failure to realize that markedness promotion and faithfulness demotion are, in reality, asymmetrical actions in the design of an OT analysis. When a markedness constraint is promoted, its possible repairs are determined by the relative ranking of the faithfulness constraints it dominates, which are always finite in number. There are only so many features of an input segment to which a surface candidate may be (un)faithful, thus limiting the number of possible FAITH constraints in a grammar. When a faithfulness constraint is demoted, on the other hand, there is no inherent control on the number of markedness constraints that may dominate it. The implication of this difference is that there is no reliable way to determine the true number of marked sequences that a given phonological process repairs. In the case of a conspiracy, the observation of multiple repairs leads to the identification of a single marked sequence, and hence a single markedness constraint motivating those repairs. With a reverse conspiracy, however, a single phonological process is observed to repair multiple sequences, and the exact subdivision of those sequences into specific markedness constraints can only be guessed at.

A possible counterargument to the insurmountability of this problem is the following: Suppose we identify a phonological process P that repairs a set of marked sequences S, where S could be subdivided into a set of two markedness constraints M_1 and M_2, a different set of two markedness constraints M_a and M_b, or a set of three markedness constraints M_i, M_{ii}, and M_{iii}. To know which subdivision is correct, we need only to inspect other areas of the grammar for the independent operation of these markedness constraints. That is, if the subdivision of S into constraints M_1 and M_2 is correct, we expect to see M_1 and M_2 independently forcing repairs in other parts of the grammar. In this way we can verify the specific set of markedness constraints used to capture the set of marked sequences S targeted by P.

As fate would have it, however, the grammar of Somali destroys the viability of this method by the fact that the alternations evident in the morphophonological environment of determiner-suffixing are unique to that environment. These processes are not present other derived environments, and many marked consonant sequences that would be repaired in the determiner-suffix environment can be found unchanged inside roots (non-derived environment blocking [NDEB]):

Examples of marked sequences that would be repaired in determiner-suffix environment	
/walaal/ ‘brother’ + /-toojo/ (‘state of being’) → [walaa**l**toojo] ‘brotherhood’	*[walaa**ʃ**oojo]
/ga**l**t**i**/ ‘foreigner, barbarian’	*[ga**ʃ**i]
/madaħ/ ‘chief’ + /-tinnimo/ (‘-ness/-ship’) → [madaħ**t**innimo] ‘leadership’	*[madaħ**d**innimo]
/waħ**t**ar/ ‘usefulness, benefit’	*[waħ**d**ar]
/h**ata**ɢ/ ‘ditch, hole in the ground’	*[ha**d**aɢ]
/raɢ**t**i/ ‘insect species with painful sting’	*[raɢ**d**i]

/afr**ika**/ 'Africa'	*[afri**g**a]
/**eka**an/ 'similarity, likeness'	*[e**ɦ**aan], *[a**ɦ**aan]
/fa**ʕk**allo/ 'anti-social behavior'	*[faʕallo]

Since these marked sequences seem to be left unrepaired in all other areas of the grammar, we cannot utilize the proposed method of inspecting for the activity of our markedness constraints outside of the determiner-suffix environment for which they were devised.

It is interesting to note, however, that although other derived environments and roots allow these marked sequences to surface unchanged (as illustrated in the examples above), a cursory dictionary search shows them to be far outnumbered by words containing "repaired versions" of the marked sequences (www.somalism.com/dictionary). For example, a search for [ɦt] yielded 17 words containing this sequence, which would be considered marked in the determiner-suffix environment and made to surface as [ɦd]. A search for the unmarked sequence [ɦd] yielded 40 words, showing that although marked sequences are allowed to surface in roots, there does exist some sort of static preference for their unmarked counterparts. This static preference is expressed actively in the alternations of the determiner-suffix environment.

With these active phonological alternations still confined to a specific morphological environment, however, the same problem remains: However we posit and subdivide a set of marked sequences into markedness constraints to account for the reverse conspiracies, we cannot verify the design by examining another section of the grammar, simply because those constraints are not active in any other environment. This is a limitation of classic OT in accounting for reverse conspiracies that are unique to a specific environment. In a roundabout way, therefore, I believe this is an argument for a cophonological approach (cf. Inkelas & Zoll 2007) to the determiner-suffix alternations in Somali. If a phonological alternation is confined to a specific morphological environment and its markedness constraints are not active anywhere else in the grammar, then the logical solution is to develop an analysis in which those markedness constraints are, by definition, ranked high only in the given morphological environment. This approach also eliminates the problem of determining the exact subdivision of markedness constraints: since their efficacy is confined to a specific morphological environment, they need only to work well enough to capture the alternations in that environment.

To summarize, classic OT is equipped to handle conspiracies, in which a single marked sequence is targeted by multiple repair strategies. The Somali data, however, shows reverse conspiracies, in which a single phonological process is used to repair a number of marked sequences. Due to the asymmetry of promoting a markedness constraint to dominate a finite number of faithfulness constraints (capturing a conspiracy) and demoting a faithfulness constraint to be dominated by any number of markedness constraints (capturing a reverse conspiracy), we cannot be certain that the markedness constraints we posit for a reverse conspiracy are indeed correct. Since the phonological alternations evident in the Somali determiner-suffix environment are found only in that morphological environment, there is no way to examine other areas of the grammar for the activity of our posited markedness constraints. The solution, therefore, is to adopt a cophonological approach and confine the activity of the relevant

markedness constraints by definition to the determiner-suffix environment. This takes the form of developing a different ranking of OT constraints (a *cophonology*) for each morphological construction. The markedness constraints relevant to the determiner-suffix alternations are ranked high only in the cophonology associated with that construction, and are dominated enough in all other cophonologies to be inactive in those environments.

5 Analysis: Describing the Somali Morphophonology (Alternation Set #1)

5.1 Brute-force approach

Having committed ourselves to an OT cophonological model, we will now examine the phonological alternations of the determiner-suffix consonants in detail to discover the markedness and faithfulness constraints active in this environment. It is important to note that it is not only the case that these alternations only appear when a determiner is suffixed to a noun, but also that they do not surface in the root even when a determiner is suffixed. This is evidence of another type of NDEB, illustrated with the following forms:

/ʃooki/ 'police station' (masc.) + /-kii/ (remote def.) → [ʃookigii] 'the police station' *[ʃoo**g**igii]
/haraati/ 'kick' (fem.) + /-tii/ (remote def.) → [haraatidii] 'the kick' *[hara**d**idii]

In each of these cases, the voicing process of /k/→[g] or /t/→[d] is motivated by a possible *VC̥V constraint that applies *only* to the initial consonant of the determiner-suffix, despite the fact that the same marked sequence is also found in the root to which the suffix is attaching. This observation necessitates a set of ROOT-FAITH constraints in the determiner-suffix cophonology to prevent any changes in the root segments.[2]

We now turn to the phonological alternations for suffix-initial /k/, which alternates among [k], [g], [ɦ], & Ø, and suffix-initial /t/, which alternates among [t], [d], [ɖ], & [ʃ] (when coalescing with stem-final /l/). The sequence /kV/ surfaces as [gV] when followed by a stem-final [g], glide, or vowel (with the exception of [e], [o], and [a]). Likewise, the sequence /tV/ surfaces as [dV] when followed by a stem-final [d], glide, or vowel. This suggests the demotion of IDENT-IO(CVoice) below the general body of FAITH to be violated as a repair for the markedness constraints *BP and *VC̥V (where BP = sequence of voiced-voiceless plosives with same POA, C̥ = voiceless consonant, V = glide/vowel):

[2] Note that the /lt/→[ʃ] process appears to defy this generalization by altering a root segment, a problem that is addressed in §5.2.

*BP, *VC̥V, FAITH ≫ IDENT-IO(CVoice)

results: /g-k/→[gg] /V-kV/→[VgV]

/d-t/→[dd] /V-tV/→[VdV]

(a dash '-' represents the morpheme boundary)

When preceded by stem-final [e], [o], or [a] (the set of non-high short vowels), however, a suffix-initial /k/ is repaired by voicing, spirantizing, and debuccalizing to [ɦi],[3] but a suffix-initial /t/ in that same sequence simply voices to [d]. To account for this disparate behavior, we propose a markedness constraint *$V_{short, -high}C_{velar}V$ (hereafter *$\nu C_{velar}V$ where ν = non-high short vowel) that militates against an intervocalic velar consonant with a preceding vowel that is non-high and short. This is added to the constraint ranking in conjunction with a demotion of the faithfulness constraints MAX(place) and IDENT-IO(cont) to make debuccalization and spirantization viable repairs:

***$\nu C_{velar}V$**, *BP, *VC̥V, FAITH ≫ IDENT-IO(CVoice), **MAX(place), IDENT-IO(cont)**

The above ranking is incomplete, however, since it allows a violation of *BP to be repaired by a single process of voicing, spirantization, or debuccalization. That is, a /g-k/ sequence could be equally repaired by /k/→[g], /k/→[ʔ], or /k/→[x], and a /d-t/ sequence could be equally repaired by /t/→[d], /t/→[ʔ], or /t/→[θ][4]. One solution to this problem is to demote IDENT-IO(CVoice) below MAX(place) and IDENT-IO(cont), making the voicing process of /k/→[g] & /t/→[d] the absolute cheapest repair. A violation of *BP then finds its cheapest repair in voicing the second consonant.

This demotion of IDENT-IO(CVoice) does not disrupt the repair for *$\nu C_{velar}V$. The combination of voicing, spirantizing, & debuccalizing of /k/→[ɦi] required to repair a [νkV] sequence is still forced by the following facts:

a) voicing /k/→[g] fails to satisfy *$\nu C_{velar}V$
b) spirantizing /k/→[x] fails to satisfy *$\nu C_{velar}V$ or *VC̥V
c) debuccalizing /k/→[ʔ] fails to satisfy *VC̥V
d) voicing and spirantizing /k/→[ɣ] fails to satisfy *$\nu C_{velar}V$
e) voicing and debuccalizing /k/ yields a voiced glottal stop, which is not a possible segment
f) spirantizing and debuccalizing /k/→[h] fails to satisfy *VC̥V
g) voicing, spirantizing, and debuccalizing /k/→[ɦi] is the only repair that satisfies both *VC̥V and *$\nu C_{velar}V$

[3] For the sake of simplicity and sanity, we will ignore the opacity produced by the trans-laryngeal assimilation of the stem-final vowels [e] and [o] to the suffix-vowel, as in /bare/ 'teacher' + /-kii/ (remote def.) → [bar**i**ɦiii] 'the teacher (remote).'

[4] Recall that Somali [t] and [d] are dental (§2).

Our constraint ranking is therefore as follows:

*$\nu C_{velar}V$, *BP, *VC̥V, FAITH ≫ IDENT-IO(cont), MAX(place) ≫ **IDENT-IO(CVoice)**

results: /g-k/→[gg] /V-kV/→[VgV] /*ν*-kV/→[*ν*fiV]
/d-t/→[dd] /V-tV/→[VdV]

This is illustrated with the following tableaux:

Full demotion of IDENT-IO(CVoice) makes /k/→[g] & /t/→[d] the cheapest repair for *BP

/g-k/	*$\nu C_{velar}V$	*BP	*VC̥V	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. gk		*!					
b. gx					*!		
c. ☞ gg							*
d. gɣ					*!		*
e. gʔ						*!	
f. gh					*!	*	
g. gɦ					*!	*	*

/d-t/	*$\nu C_{velar}V$	*BP	*VC̥V	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. dt		*!					
b. dθ					*!		
c. ☞ dd							*
d. dð					*!		*
e. dʔ						*!	
f. dh					*!	*	
g. dɦ					*!	*	*

/k/→[g] & /t/→[d] is also the cheapest repair for *VC̥V

/u-ka/	*$\nu C_{velar}V$	*BP	*VC̥V	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. uka			*!				
b. uxa			*!		*		
c. ☞ uga							*
d. uɣa					*!		*
e. uʔa			*!			*	
f. uha			*!		*	*	
g. uɦa					*!	*	*

/u-ta/	*νC$_{velar}$V	*BP	*VÇV	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. uta			*!				
b. uθa			*!		*		
c. ☞ uda							*
d. uða					*!		*
e. uʔa			*!			*	
f. uha			*!		*	*	
g. uɦa					*!	*	*

/k/→[ɦ] is still forced as the sole repair for *νC$_{velar}$V

/a-ka/	*νC$_{velar}$V	*BP	*VÇV	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. aka	*!		*				
b. axa	*!		*		*		
c. aga	*!						*
d. aɣa	*!				*		*
e. aʔa			*!			*	
f. aha			*!		*	*	
g. ☞ aɦa					*	*	*

/k/ & /t/ are allowed to surface unchanged in unmarked sequences

/r-ka/	*νC$_{velar}$V	*BP	*VÇV	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. ☞ rka							
b. rxa					*!		
c. rga							*!
d. rɣa					*!		*
e. rʔa						*!	
f. rha					*!	*	
g. rɦa					*!	*	*

/r-ta/	*νC$_{velar}$V	*BP	*VÇV	FAITH	IDENT(cont)	MAX(place)	IDENT(CVoice)
a. ☞ rta							
b. rθa					*!		
c. rda							*!
d. rða					*!		*
e. rʔa						*!	
f. rha					*!	*	
g. rɦa					*!	*	*

Another alternation involves the behavior of /k/ & /t/ when preceded by one of the stem-final consonants [ɢ], [χ], [ʕ], [ħ], [ʔ], and [h]: /k/ deletes, but /t/ voices to [d]. To capture the unity of environments that cause this alternation, we must weaken our working assumption of [h] & [ʔ] as "placeless" consonants to include the place-specification of "laryngeal." If [h] & [ʔ] are considered to be both "placeless" and "laryngeal," then we can describe the set of consonants that trigger this alternation as "post-velar" (i.e. uvular, pharyngeal, or laryngeal). Capturing the alternations of /k/ & /t/ in our constraint ranking is then accomplished by:

a) positing the markedness constraint *$C_{post\text{-}velar}$C̥ (hereafter *C_{pv}C̥), which militates against a post-velar consonant followed by a voiceless consonant (forcing /t/→[d])
b) positing the markedness constraint *$C_{post\text{-}velar}C_{post\text{-}palatal}$ (hereafter *$C_{pv}C_{pp}$), which militates against any sequence of a post-velar and post-palatal consonant (preventing a repair that voices /k/ to [g] or voices & debuccalizes /k/ to the laryngeal [ɦ])
c) demoting MAX-C (forcing /k/→Ø) from the general body of FAITH to a position just above the other demoted faithfulness constraints (a position that precludes the use of /k/-deletion as a repair for the other marked sequences).

***C_{pv}C̥, *$C_{pv}C_{pp}$**, **ν*C_{velar}V, *BP, *VC̥V, FAITH ≫ **MAX-C** ≫ IDENT-IO(cont), MAX(place) ≫ IDENT-IO(CVoice)

results: /C_{pv}-t/→[C_{pv}d] /C_{pv}-k/→[C_{pv}]

e.g. /ʔ-t/→[ʔd] /ʔ-k/→[ʔ]

Voicing of /t/→[d] is the cheapest repair for *C_{pv}C̥

/ʔ-t/	*C_{pv}C̥	*$C_{pv}C_{pp}$	*$ν$$C_{velar}$V	*BP	*VC̥V	FAITH	MAX-C	IDENT(cont)	MAX(plc)	IDENT(CVoice)
a. ʔt	*!									
b. ʔθ	*!							*		
c. ☞ ʔd										*
d. ʔð								*!		*
e. ʔʔ	*!	*							*	
f. ʔh	*!	*						*	*	
g. ʔɦ		*!						*	*	*
h. ʔ_							*!			

Deletion of /k/→Ø is forced as the sole repair for *$C_{pv}\underset{\circ}{C}$ and *$C_{pv}C_{pp}$

/ʔ-k/	*$C_{pv}\underset{\circ}{C}$	*$C_{pv}C_{pp}$	*$\nu C_{velar}V$	*BP	*$V\underset{\circ}{C}V$	FAITH	MAX-C	IDENT(cont)	MAX(plc)	IDENT(CVoice)
a. ʔk	*!	*								
b. ʔx	*!	*						*		
c. ʔg		*!								*
d. ʔɣ		*!						*		*
e. ʔʔ	*!	*							*	
f. ʔh	*!	*						*	*	
g. ʔɦ		*!						*	*	*
h. ☞ ʔ_							*			

At this point, all the alternations of suffix-initial /k/ are captured by our constraint ranking. The two remaining processes involve suffix-initial /t/:

a) the anomalous coalescence of stem-final /l/ and suffix-initial /t/ to yield [ʃ]
b) the progressive retroflex assimilation of /ɖ-t/→[ɖɖ]

Under our current framework, the best solution for describing the /l-t/→[ʃ] process is to rank a stipulative constraint such as INPUT(lt)≈Output(ʃ) above ROOT-FAITH, allowing the root-final segment /l/ to coalesce with suffix-initial /t/. Finally, if we assume that /ɖ/ and /t/ share the same general place of articulation [CORONAL], the cheapest repair for the /t/ in a marked /ɖ-t/ sequence is to be voiced to [d]. To yield [ɖ], then, we posit the markedness constraint *$R_{cor}C_{cor}$, which militates against a sequence of coronal consonants in which the first is [+retroflex][5] and the second is [–retroflex], and also demote IDENT-IO(retroflex) a level equal with IDENT-IO(voice).[6] This forces the /t/ in a marked /ɖ-t/ sequence to be repaired by voicing and retroflexing to [ɖ]:

INPUT(lt)≈Output(ʃ) ≫ ***$\mathbf{R_{cor}C_{cor}}$**, *$C_{pv}\underset{\circ}{C}$, *$C_{pv}C_{pp}$, *$\nu C_{velar}V$, *BP, *$V\underset{\circ}{C}V$, **ROOT-FAITH**, FAITH ≫ MAX-C ≫ IDENT-IO(cont), MAX(place) ≫ IDENT-IO(CVoice), **IDENT-IO(retroflex)**

results: /l-t/→[ʃ] /ɖ-t/→[ɖɖ]

[5] [+retroflex] can be considered an abstract representation of the group of features that a define retroflex consonant.

[6] IDENT-IO(retroflex) can safely occupy this low position without disrupting our analysis up to this point, simply because it fails to provide an avenue of repair for any of the other marked sequences already described.

/t/→[ɖ] is the cheapest repair for *$R_{cor}C_{cor}$ via low-ranked IDENT-IO(retroflex) and IDENT-IO(CVoice)
(The inactive markedness constraints *$C_{pv}C_{\underset{\circ}{}}$, *$C_{pv}C_{pp}$, *$\nu C_{velar}V$, and *$V\underset{\circ}{C}V$ are omitted from the following tableau for the sake of space.)

/ɖ-t/	*$R_{cor}C_{cor}$	*BP	FAITH	MAX-C	IDENT(cont)	MAX(place)	IDENT(CVoice)	IDENT(retroflex)
a. ɖt	*!	*						
b. ɖd	*!						*	
c. ☞ ɖɖ							*	*
d. ɖz̪					*!		*	*
e. ɖt̪		*!						*
f. ɖs̪					*!			*
g. ɖ_				*!				
h. ɖʔ						*!		
i. ɖh					*!	*		
j. dɦ					*!	*	*	

Although this analysis ostensibly works, it is rather cumbersome and unwieldy. As such, we will now sketch another approach that seeks unity among the various repairs for /k/ and /t/.

5.2 Scales approach

When ordered as follows, the surface forms of determiner-suffix /k/ convey a sequence of lenition that proceeds from most-faithful to least-faithful:

position:	0	1	2	3
form:	[k]	[g]	[ɦ]	Ø
faithfulness violations:	(none)	IDENT(CVoice) *IDENT(cont)*[7]	IDENT(CVoice) IDENT(cont) MAX(place)	MAX-C (essentially all IDENT constraints)

This scale captures the limited number of repairs available for marked sequences in the determiner-suffix environment. Notice that such segments as [ʔ] and [h] are not listed, since they never surface as variants of underlying /k/. When a marked /k/ sequence is present in the input, it may only be repaired by moving to a position on the scale. In addition, the "worse" a marked sequence is, the further along the scale its first possible repair will be. For example, if a certain marked sequence is repaired by position #2, then positions #0 and #1 were, by stipulation, not possible surface forms for /k/ in that sequence (in this cophonology).

[7] Armstrong (1964) observes that Somali /g/ sometimes appears as [ɣ] between vowels, which suggests that the IDENT(cont) violation of position #2 may already be a part of position #1.

Potential overlap among marked environments, then, is not a problem for this analysis, since a marked sequence proceeds along the scale in search of a repair and stops at the first possible position. For instance, a [aka] sequence finds its first possible repair in position #2, surfacing as [a**ɦ**a]. A [uka] sequence, though it shares a violation of *VC̥V with [aka], finds its first possible repair as early as position #1, surfacing as [u**g**a]. It gains nothing by moving to position #3 and surfacing as [u**ɦ**a].

Before describing the marked sequences that require the repairs at each position, it should be noted that it is possible to reanalyze position #3 as an empty C slot, akin to deleting a root node. The only marked sequences that make use of the repair in position #3 are those in which /k/ is preceded by one of the post-velar consonants [ɢ], [χ], [ʕ], [ħ], [ʔ], and [h], which never appear as geminates. The analysis of a position #3 repair, then, would be that the root node of /k/ is deleted, leaving an empty C slot to which the preceding consonant spreads. Since post-velar (and velar) geminates are not allowed in the language, this illegitimate geminate is reduced to a single consonant, yielding the same result if /k/ had simply deleted. For the sake of simplicity, we will leave position #3 as Ø, but this C-slot analysis will be important for our investigation of the /t/-alternations.

The marked sequences that correspond to each /k/-repair can now be listed at their respective positions:

position:	0	1	2	3
form:	[k]	[g]	[ɦ]	Ø
faithfulness violations:	(none)	IDENT(CVoice) *IDENT(cont)*	IDENT(CVoice) IDENT(cont) MAX(place)	MAX-C (essentially all IDENT constraints)
marked sequences:	(unmarked)	*VC̥V, *BP	*$\nu C_{velar} V$	*$C_{post\text{-}velar} C_{post\text{-}palatal}$
examples:	/b-k/→[bk]	/g-k/→[gg] /u-ka/→[uga]	/a-ka/→[aɦa]	/ʔ-k/→[ʔ]

An examination of the surface forms of determiner-suffix /t/ yields a similar scale in which the repairs to underlying /t/ proceed from most-faithful to least-faithful along a path of lenition:

position:	0	1	2
form:	[t]	[d]	C
faithfulness violations:	(none)	IDENT(CVoice) *IDENT(cont)*[8]	deletion of root node
marked sequences:	(unmarked)	*VC̥V, *BP *$C_{post\text{-}velar} C̥$	*lt *$R_{cor} C_{cor}$
examples:	/b-t/→[bt]	/a-ta/→[ada] /d-t/→[dd] /ʔ-t/→[ʔd]	/l-t/→[ʃ] /ɖ-t/→[ɖɖ]

[8] As with /g/, Armstrong (1964) observes that Somali /d/ often lenites to [ð] between vowels.

The fully-faithful form [t] in position #0 surfaces when the suffix-initial /t/ is part of an unmarked sequence. The repair in position #1, /t/→[d], is preferred when the offending /t/ is intervocalic (*VC̥V), follows a [d] (*BP), or follows a post-velar consonant (*$C_{\text{post-velar}}$C̥)

Finally, the most drastic repair is in position #2 and involves the deletion of /t/'s root node, leaving an empty C slot. The preceding consonant (/ɖ/ or /l/) then spreads its features to fill that C. In the case of /ɖ/, a geminate [ɖɖ] is formed. In the case of /l/, the segment that surfaces is [ʃ].

The reasoning behind the surfacing of [ʃ] when an /l/ spreads is as follows: According to Armstrong (1964), a word-final single /l/ is often partially voiceless in Somali (e.g. [mee$\widehat{\text{l̥l̥}}$] 'stick'). When a /t/-initial determiner suffix is attached to a /l/-final word, the root node of the /t/ is deleted to repair the marked sequence *lt, leaving behind an empty C slot (e.g. /mee$\widehat{\text{l̥l̥}}$-Ca/). The partially voiceless [$\widehat{\text{l̥l̥}}$] spreads rightward to the C slot, forming [$\widehat{\text{l̥l̥l̥}}$], a geminate consonant that is not present in the inventory of Somali. Since a geminate voiceless /l/ sounds much like [ʃ], an actual Somali segment, the [$\widehat{\text{l̥l̥l̥}}$] consonant is perceived as [ʃ] (e.g. [meeʃa] 'the stick'). This approach could be considered a diachronic as well as synchronic explanation, and also has the benefit of conforming to the principle of ROOT-FAITH presented in §5.1 to account for non-derived environment blocking: because the repair in position #2 simply deletes the root node of suffix-initial /t/, it does not violate ROOT-FAITH. Neither does the spreading of /l/ to the empty C slot violate ROOT-FAITH. Only in phonetic production and/or perception is the resultant [$\widehat{\text{l̥l̥l̥}}$] consonant reinterpreted as [ʃ].

The approach developed in this section, which accounts for the determiner-suffix /k/ and /t/ alternations by means of two scales, is much more intuitive than the previous brute-force method that resulted in an unwieldy set of constraints. For the purposes of this paper, however, a full OT implementation of the scale approach cannot be presented here. Let us assume, for the sake of argument, that we have a set of two fully-developed cophonologies to account for the root-internal and determiner-suffix phonological processes, as diagrammed here (modeled after Yu 2000):

$$\varphi_{\text{determiner suffix1}}([\text{NOUN}], /\text{DET SUFFIX}/)$$
$$\Big|$$
$$\varphi_{\text{root}}(/\text{NOUN}/)$$

example:

$$\varphi_{\text{determiner suffix1}}([\text{gabaɖ}], /\text{-ta}/) = [\text{gabaɖɖa}] \text{ 'the girl'}$$
$$\Big|$$
$$\varphi_{\text{root}}(/\text{gabɖ}/) = [\text{gabaɖ}] \text{ 'girl' (fem.)}$$

Note that the $\varphi_{\text{determiner suffix1}}$ cophonology contains a set of ROOT-FAITH constraints to prevent its repairs from acting on marked sequences contained inside the root [NOUN].

This cophonological setup will aid us in accounting for another property of the Somali determiner-suffix alternations, presented in the next section.

6 Data: Alternation Set #2

As described in §3, Somali nouns can combine with a number of different determiner suffixes: definite articles, demonstratives, possessives, and interrogatives. When a noun is suffixed with a single possessive determiner, whose initial consonant (/k/ or /t/) undergoes the alternations described, the semantic interpretation is one of inalienable possession:

/lug/ 'leg' (fem.) + /-teed/ (3sg fem. possessive) → [lugteed] 'her leg'
/aabbe/ 'father' (masc.) + /-kaj/ (1sg possessive) → [aabbaɦaj] 'my father'

The complete set of possessive suffixes is given here for reference (relisted from §3):

Possessives

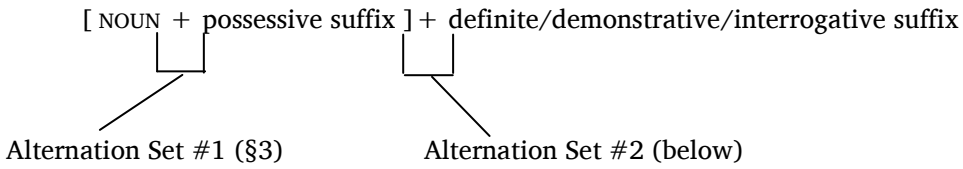
-kaj/taj	1sg 'my'	-kaja/taja	1pl exclusive 'our'
-kaa/taa	2sg 'your'	-keen/teen	1pl inclusive 'our'
-kiis/tiis	3sg masc 'his'	-kiin/tiin	2pl 'your'
-keed/teed	3sg fem 'her'	-kood/tood	3pl 'their'

To form the alienable possessive, another determiner is suffixed onto the end of the noun + possessive stem. This second suffix (shown below in boldface) can be any from the full range of determiners (with the exception of another possessive):

/guri/ 'house' (masc.) + /-kaj/ (1sg poss.) + /**-ka**/ (non-remote def.) → [gurigaj**ga**] 'my house'
/faras/ 'horse' (masc.) + /-kaa/ (2sg poss.) + /**-koo**/ (far dem.) → [faraskaa**goo**] 'that (far) horse of yours'
/ul/ 'stick' (fem.) + /-taja/ (1pl excl. poss.) + /**-tee**/ (interrog.) → [uʃaja**dee**] 'which stick of ours?'

While the possessive determiner suffixed directly to the noun undergoes the alternations described in §3, the second determiner suffix exhibits a totally different set of alternations, listed below.

Alienable possessive formation in Somali



Surface forms of suffix-initial /k/, /t/ when attached to noun + possessive stem

[g], [d] after stem-final [j], [a]

e.g. /guri/ 'house' (masc.) + /-kaa/ (2sg poss.)
→ [gurigaa] + /-ka/ (non-remote def.) Alternation #1
→ [gurigaa**g**a] 'your house' Alternation #2

/ul/ 'stick' (fem.) + /-taj/ (1sg poss.)
→ [uʃaj] + /-tan/ (near demons.) Alternation #1
→ [uʃaj**d**an] 'this stick of mine' Alternation #2

[n] after stem-final [n]

e.g. /faras/ 'horse' (masc.) + /-kiin/ (2pl poss.)
→ [faraskiin] + /-kee/ (interrog.) Alternation #1
→ [faraskiin**n**ee] 'which horse of yours?' Alternation #2

/hees/ 'song' (fem.) + /-teen/ (1pl inclusive poss.)
→ [heesteen] + /-tii/ (remote def.) Alternation #1
→ [heesteen**n**ii] 'our song (remote)' Alternation #2

Ø after stem-final [d], [s]

e.g. /ɖuu/ 'leather garment' (masc.) + /-kiis/ (3sg masc. poss.)
→ [ɖuugiis] + /-kan/ (near demons.) Alternation #1
→ [ɖuugiis_an] 'this leather garment of his' Alternation #2

/loʔ/ 'cattle' (fem.) + /-tood/ (3pl poss.)
→ [loʔdood] + /-teer/ (middle distance demons.) Alternation #1
→ [loʔdood_eer] 'those cattle of theirs' (mid. dist.) Alternation #2

Before accounting for the difference in behavior of the determiner morphemes when suffixed in different morphological positions, we can analyze this second set of phonological alternations with a scales approach, presented in the next section.

7 Analysis: Describing Alternation Set #2

As with the first set, the second set of /k/ & /t/ alternations submits to an intuitive scales approach in which the repairs are ordered from most-faithful to least-faithful. This set, however, utilizes a single scale of lenition for both consonants:

position:	1	2	3
form:	[g] / [d]	C	Ø
faithfulness violations:	IDENT(CVoice) *IDENT(cont)*[9]	deletion of root node	MAX-C (essentially all IDENT constraints)
marked sequences:	*VC̥V	*nP (P = voiceless plosive)	*CP
examples:	/a-ka/→[aga] /a-ta/→[ada]	/n-k/→[nn] /n-t/→[nn]	/s-k/→[s] /d-t/→[d]

/k/ and /t/ are never allowed to surface unchanged as [k] and [t] in this morphological environment, and thus position #0 is not present on the scale. To amend a marked *VC̥V sequence, the repair in position #1 must be enacted by leniting the offending intervocalic[10] consonant to a voiced one. When *nP is violated by a sequence in which /n/ precedes /k/ or /t/, the root node of the plosive is deleted (position #2). This leaves an empty C slot to which /n/ spreads, forming the geminate [nn]. Finally, when /k/ or /t/ is preceded by any other consonant in this morphological environment (essentially [d] or [s]), the entire suffix-initial consonant is deleted (position #3).

In addition to an OT implementation of the above scale, this cophonology also includes a set of undominated STEM-FAITH constraints that prevents its repairs from affecting marked sequences contained inside the stem. This approach isolates each set of alternations (described by scales) to its respective domain in the morphologically-complex word.

8 A Unified Model: Scales & Cophonologies

Having developed a cophonology for each set of suffix-initial consonant alternations, we finally combine them to form the whole "machine" that outputs a Somali word:

$$\varphi_{\text{determiner suffix2}}([\text{NOUN} + \text{POSS SUFFIX}], /\text{DET SUFFIX}/)$$
$$\varphi_{\text{determiner suffix1}}([\text{NOUN}], /\text{POSS SUFFIX}/)$$
$$\varphi_{\text{root}}(/\text{NOUN}/)$$

where φ_{root} = cophonology describing general Somali root phonology (not discussed in this paper)
$\varphi_{\text{determiner suffix1}}$ = cophonology with OT scales implementation of Alternation Set #1 and undominated ROOT-FAITH (§5.2)
$\varphi_{\text{determiner suffix2}}$ = cophonology with OT scales implementation of Alternation Set #2 and undominated STEM-FAITH (§7)

[9] Recall that /g/ and /d/ can lenite to [ɣ] and [ð] between vowels.

[10] "Intervocalic" here also refers to post-glide, prevocalic positions such as [j**k**a].

Examples

$\varphi_{\text{determiner suffix2}}$([gurigeen], /-kan/) = [gurigeen**n**an] ‘this house of ours (inc.)’ (*house-our-this*)

$\varphi_{\text{determiner suffix1}}$([guri], /-keen/) = [guri**g**een] (*house-our-*)

φ_{root}(/guri/) = [guri] ‘house’ (masc.)

$\varphi_{\text{determiner suffix2}}$([uʃiis], /-tee/) = [uʃiis_ee] ‘which stick of his?’ (*stick-his-which*)

$\varphi_{\text{determiner suffix1}}$([ul], /-tiis/) = [u**ʃ**iis] (*stick-his-*)

φ_{root}(/ul/) = [ul] ‘stick’ (fem.)

$\varphi_{\text{determiner suffix2}}$([saʕaj], /-koo/) = [saʕaj**g**oo] ‘that (far away) cow of mine’ (*cow-my-that*)

$\varphi_{\text{determiner suffix1}}$([saʕ], /-kaj/) = [saʕ_aj] (*cow-my-*)

φ_{root}(/saʕ/) = [saʕ] ‘cow’ (masc.)

$\varphi_{\text{determiner suffix2}}$([gabaɖɖood], /-tii/) = [gabaɖɖood_ii] ‘their girl (remote)’ (*girl-their-the*)

$\varphi_{\text{determiner suffix1}}$([gabaɖ], /-tood/) = [gabaɖ**ɖ**ood] (*girl-their-*)

φ_{root}(/gabɖ/) = [gabaɖ] ‘girl’ (fem.)

$\varphi_{\text{determiner suffix2}}$([faraskaa], /-ka/) = [faraskaa**g**a] ‘your (sg.) horse’ (*horse-your-the*)

$\varphi_{\text{determiner suffix1}}$([faras], /-kaa/) = [fara**s**kaa] (*horse-your-*)

φ_{root}(/fars/) = [faras] ‘horse’ (masc.)

9 Conclusion

An examination of the morphophonological alternations exhibited by Somali determiner-suffixes has shown them to submit most easily to a cophonological model in which each cophonology implements its respective consonant alternations by means of an OT-based scale of repair strategies. Although this approach was not fleshed out to the fullest extent in this paper, significant groundwork was laid.

Future research could also extend this cophonological approach to account for some interesting morphophonological phenomena not discussed in the body of this paper. When a noun or other such stem is suffixed with a determiner morpheme, the stem undergoes a process of "Accent Shift" in which its tones are altered in a predictable pattern. This pattern of shift is specific to each determiner morpheme. For example, the interrogative determiners /-kee/ and /-tee/ lower all high tones in the stem, while the remote definite determiners /-kii/ and /-tii/ simply append their own high tones. These processes in Somali, which call to mind the cyclic stress assignment patterns in Tohono O'odham described by Yu (2000) with a cophonological approach, are ripe ground for an extension of the cophonological model developed here.

References

ARMSTRONG, LILIAS. 1964. *The phonetic structure of Somali.* Berlin: Gregg Press.

INKELAS, SHARON, and ZOLL, CHERYL. 2007. Is grammar dependence real? A comparison between cophonological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45.133-171.

KAGER, RENÉ. 1999. *Optimality theory*. Cambridge: Cambridge University Press.

KISSEBERTH, CHARLES. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1.291-306.

SAEED, JOHN. 1999. *Somali.* Philadelphia: John Benjamins Publishing Co.

SOMALISM.COM. 2005-. *Somali-English dictionary*. <http://www.somalism.com/dictionary/>.

YU, ALAN. 2000. *Stress assignment in Tohono O'odham.* Phonology 17.117-135.

ZORC, DAVID, and OSMAN, MADINA. 1993. *Somali-English dictionary with English index.* 3rd edition. Kensington: Dunwoody Press.