ryan.txanson@gmail.com

Somali Vowel Harmony

HIF-3021: Candidate 5

June 4, 2010

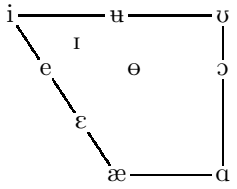
1 Introduction

Somali is a Cushitic language spoken on the horn of Africa by approximately 16 million people. It is an agglutinative and tonal language, and is host to a unique kind of vowel harmony in which the domain of harmony appears to be larger than the word, and the domain often ends up being as large as whole phrases of words, and is clearly sensitive to syntactic boundaries. Little work has been done to explain the details of how the phenomenon works, outside of the initial descriptive work on the phenomenon (Andrzejewski, 1955). This paper will explore the behaviour of this phenomenon, and model of it in Optimality Theory: using mainly prosodic wellformedness constraints (Selkirk, 1995), and generalized alignment constraints (McCarthy & Prince, 1993). Although the phenomenon seems rather complex and variable, it seems that with the aid of OT, it is easily able to be modelled.

1.1 Somali vowel system

Somali has 10 basic vowels, which fall into two groups: advanced tongue root [+ATR] and [-ATR]. The vowels occur long and short, and there are diphthongs which both have +ATR and -ATR versions. The [+ATR] vowel set also contains some quality differences: they are all much further front than their [-ATR] counterparts. Likely a historical change, since [+ATR] vowels are more likely to be pushed front. (Krämer, 2003)

(1) Somali vowels and their approximate positions (Liberman, 2002):



The above chart shows all of the vowels, however in this paper the symbol [É̱] will be used to represent the +ATR back, mid vowel. In some of the following examples, underlining is be used to draw more attention to +ATR spans.

2 ATR Harmony in Somali

A review of the data available in the initial survey of this phenomenon (Andrzejewski, 1955) reveals some facts which are explicitly laid out by the author, and some which may be found on further analysis of the data. Vowel harmony in Somali appears to be both a regressive and progressive process, in which +ATR spreads further to the left than to the right. The process of spreading may cause vowels which otherwise surface as -ATR to appear as +ATR if they fall to the left of an underlyingly +ATR segment. -ATR vowels are not able to spread over +ATR vowels, which must always surface.

Individual words pronounced on their own have vowels which are all +ATR or -ATR:

(2)

[+ATR]		[-ATR]	
mægæːlə	'town'	fanɖaːl	'spoon'
ʉgæːs	'chieftain'	bɔqɔr	'chieftain'
biyə	'water'	ʜɔːlɔ	'livestock, wealth'
berberæ	'a town name'	bʊrʕɔ	'a town name'

The language does allow for minimal pairs of both short (3) and long vowels (4), but there appears to be a strong correspondence between +ATR/-ATR and specific morphology: most of the -ATR words represent imperatives while most of the +ATR words here represent 3PM.PAST, with a few exceptions:

(3)

[+ATR]		[-ATR]	
ɖis	'build.3PM.PAST'	ɖɪs	'build.IMPER.SG'
hel	'find.3PM.PAST'	hɛl	'find.IMPER.SG'
kæb	'set fracture.3PM.PAST'	kɑb	'sandal'
qəd	'dig.3PM.PAST'	qɔd	'dig.IMPER.SG'
tʉs	'show.3PM.PAST'	tʊs	'show.IMPER.SG'

(4)

[+ATR]		[-ATR]	
diːdæj	‘faint.1PSG.PAST’	dɪːdɑj	‘refuse.1PSG.PAST’
heːs	‘sing.3PM.PAST’	hɛːs	‘a type of song/poem’
læːb	‘fold.3PM.PAST’	lɑb	‘chest (thorax)’
ɵːd	‘enclose.3PM.PAST’	ɔːd	‘thorny fence’
dʉːʃæj	‘attack.3PFEM.PAST’	dʊːʃɑj	‘fly.3PFEM.PAST’

(Andrzejewski, 1955)

It is also clear from the data that some suffixes have underlying +ATR specifications and can regressively affect the word they are attached to (5), and +ATR words may progressively spread their +ATR feature to the suffixes attached to them (6).

(5)

wɑdːɔ	‘road’	wædːɵ-jin	‘road.PL’
ɛːg	‘look’	eːg-æjæj	‘look.PASTPROG’

(6)

wæːj	‘fail’	wæːjæj	‘fail.PAST’
ɖɑqɑːq	‘go away’	ɖɑqɑːq-ɑj	‘go away.PAST’

Compound words show no progressive spreading through the other elements of the compounds, but do show regressive spreading, which points at the domain of rightward spreading as ending at the prosodic word (PRWD) boundary (with one exception, to be explained).

(7)

a)	dɑb	+	qɑd	->	dɑbqɑːd
	‘fire’		‘take’		‘incense burner’
b)	rɛːr	+	mægæːlɵ	->	reːrmægæːlɵ
	‘people’		‘town’		‘townsfolk’
c)	dæːn	+	ɖɛːr	->	dæːnɖɛːr
	‘jaw’		‘long’		‘strong camels of a herd’
d)	tʃeːni	+	wɑrɛːn	->	tʃeːniwɑrɛːn
	‘front leg’		‘stab’		‘pneumonia’

Function words in Somali typically share the same +ATR specification as adjacent lexical words, although typically only function words to the left share this; and as we’ve seen +ATR spreading does not spread right over prosodic word boundaries, function words at the ends of short sentences to the right of lexical words can be spread to. The first two of the following examples (8) show that the conditional marker is progressively spread to from the infinitive verbs, which are underlyingly +ATR and -ATR. In addition, the person marker clitic which attaches to the preceding focus marker harmonizes. The remain-

ing two (9) show essentially the same thing, except in this case with a question marker (preceding the noun), and a focus marker (following).

(8) a. w-ʉː tegi læhæː
FOC-3PM go.INF COND
'He would go.'

b. w-ʊː ɪmɑn lɑhɑː
FOC-3PM come.INF COND
'He would come.'

(9) a. mæ libæːʜ bæː
Q lion FOC
'Is it a lion?'

b. mɑ ʃɑbɛːl bɑː
Q leopard FOC
'Is it a leopard?'

The pronominal clitic markers (as in (8)) do not necessarily harmonize with the word that they attach to however. Either this is because they occur outside of a prosodic word boundary, or it is because they may be even further away syntactically than just clitics. The following examples show this, the first (10) shows a sentence in which the focus marker stands on its own, but the following (10b), shows the focus marker and personal clitic fused to the preceding focused word. This results in a word in which the personal clitic is directly attached to the focused noun, although it is clear that there are additional syntactic boundaries underlyingly.

(10) a. berberæ bʊː qɑb-tɑj
berberæ FOC-3PMASC go-3PM.PAST
'He went to Berbera.'

b. berber=ʊː qɑb-tɑj
berberæ=FOC-3PMASC go-3PM.PAST
'He went to Berbera.'

In his description of Somali vowel harmony, Andrzejewski noted that things such as speech rate and pauses (11) could delimit the "harmonic groups". With a faster speech rate, there is more spreading. Following is also a list the underlying representations (12) of the words as provided or otherwise discovered by analysis of the data in (Andrzejewski, 1955).

(11) a. beːr-tiː mæ sɵː iːbsæ-tæj
garden-DEF Q DEIC buy-2PSG.PAST
'Have you bought the garden?'

b. bɛːr-tɪː # mæ sɵː iːbsæ-tæj
garden-DEF PAUSE Q DEIC buy-2PSG.PAST
'Have you bought the garden?'

(12)

bɛːr	'garden'
mɑ	Q
sɔː	'toward'
iːbsɵ	'buy'

In summary, we see that vowel harmony spreading is stopped to the right by prosodic word boundaries, except for in cases where the function word ends up at the end of an utterance. Spreading leftward of +ATR continues over PRWD boundaries but is stopped by pauses in speech.

3 Syntax-Phonology Interface

The previous section detailed the behaviour of +ATR harmony in Somali, and outlined what is at the core of the phenomenon: all PRWDs are either +ATR or -ATR, and the PRWD includes suffixes but not clitics. Spreading to the right past the PRWD boundary is possible when one function word falls at the end of the utterance, but otherwise, spreading leftward is mostly unimpeded, except by pauses in speech. If the spans of +ATR are contained by ϕ-phrases, then the edges of ϕ phrases may be aligned to either the right edge of the PRWD, or to the right of an utterance. Spreading is possible leftward across ϕ-phrase boundaries.

This paper adopts the prosodic hierarchy (Selkirk, 1981; Truckenbrodt, 2007), which is summarized with the following structure:

(13) **Prosodic Hierarchy**

U	Utterance
|	
IP	Intonational phrase
|	
ϕ	Phonological phrase
|	
ω	Prosodic word
|	
Σ	foot
|	
σ	syllable

3.1 Prosodification

Based on what we've seen with suffixes, the boundary for rightward spreading appears to be the PrWd boundary. The reason for this is that spreading of +ATR appears to be blocked from going rightward in compound words (14), but occurs both leftward and rightward within PrWds from either root to suffix or suffix to root (15).

(14)

b)	rɛːr	+	mægæːlə	->	[reːr]$_\omega$[mægæːlə]$_\omega$
	'people'		'town'		'townsfolk'
c)	dæːn	+	ɖɛːr	->	[dæːn]$_\omega$[ɖɛːr]$_\omega$
	'jaw'		'long'		'strong camels of a herd'

(15)

wɑdːɔ	'road'	[wædːə-jin]$_\omega$	'road.Pl'
ɛːg	'look'	[eːg-æjæj]$_\omega$	'look.PastProg'
wæːj	'fail'	[wæːj-æj]$_\omega$	'fail.Past'
ɖɑqɑːq	'go away'	[ɖɑqɑːq-ɑj]$_\omega$	'look.PastProg'

Additionally, rightward spreading does not occur to function words such as clitics (except when they are at the end of a intonational phrase), such as the pronoun [ʉː] '3pMasc', which is attached to the place name [berberæ] in the following sentence (16). Clitics as a result, must be attached outside of the prosodic word boundary; and likely belong to a following ϕ-phrase as a function word. (Selkirk, 1995)

(16) berber=ʊː qɑb-tɑj
berber=Foc.3pMasc go-3pM.Past
'He went to Berbera.'

The first constraints to model this follow, with the beginning initial ranking in (19). NonRecursivity, one of the four prosodic well-formedness constraint (Selkirk, 1995), is required because of the assumption that spreading occurs within ϕ-phrases (and leftward in intonational phrases, as will be discussed later). There is no explicit evidence for recursive ϕ-structures in the data, and so it seems best to rule them out for the sake of simplicity; but ruling them out with NonRecursivity also rules out recursive PrWd structures.

(17) Align(PrWd,R;ϕ,R): "The right edge of each PrWd is aligned with the right edge of a ϕ-phrase."

(18) NonRecursivity: No C^i dominates C^i. (No foot dominates a foot)

(19) NonRecursivity $\gg$ Align-PrWd-R

The following tableau shows the ranking for compound words, such as [reːrmægæːlə] 'townspeople', where +ATR spreads leftward; or [dæːnɖɛr] 'strong camels of a herd', where spreading does not occur

rightward because of ϕ-phrase boundaries which are aligned with the PrWds. What is important to see in this tableau is only the placement of the ϕ phrases, so +ATR features are left out.

(20)

PrWd PrWd	NonRecursivity	Align-PrWd-R
a. (PrWd PrWd)$_\varphi$		*!
b. ☞ (PrWd)$_\varphi$ (PrWd)$_\varphi$		
c. ((PrWd)$_\varphi$(PrWd)$_\varphi$)$_\varphi$	*!*	

The winning candidate (b) has two ϕ-phrases which mark the right edges of the PrWds, however since recursive ϕ-phrases are disallowed, candidate (c) is ruled out. Although the winning output represents a word with complete +ATR, and one would expect only one ϕ-phrase, the required ϕ-structure to explain a word in which there is *not* spreading (such as [[dæːn]$_\varphi$[ɖɛːr]$_\varphi$]$_I$) is that there are two ϕ-phrases. Otherwise, we must come up with some complex explanation for alignment of ϕ-phrases to +ATR vowels.

Next we see that function words including those that occur to the right of a PrWd, must also be included in ϕ-phrases because +ATR can spread to them (21). This can be modelled with another alignment constraint, because what appears to cause these function words to join preceding ϕ-phrases is that the right edge of a ϕ-phrase must be align with the right edge of an intonational phrase. This however will result in several violations of a prosodic well-formedness constraint, Exhaustivity, defined below (23). This constraint is quite low-ranked, so it will be left out of the tableaux for now since its violations of each candidate will be equivalent and do not yet result in significant differences. There will always be a few violations of it for every candidate as well; the analysis depends on ϕ-phrases dominating function words, which are simply feet or monosyllables, so each instance of a function word will result in violations.

(21) a. mæ libæːʜ bæː
Q lion Foc
'Is it a lion?'

b. ma ʃɑbɛːl baː
Q leopard Foc
'Is it a leopard?'

(22) Align(IP, R; ϕ, R) / Align-ϕ-R: "The right edge of each IP is aligned with the right edge of a ϕ-phrase."

(23) Exhaustivity: No C^j immediately dominates a constituent C^i, j < i-1, e.g. "No ϕP immediately dominates a PrWd."

(24) NonRecursivity ≫ Align-ϕ-R, Align-PrWd-R (≫ Exhaustivity)

This ranking is demonstrated in the following tableau, which shows that candidate (a) violates the PrWd alignment constraint, however it is okay since the IP alignment constraint is more highly ranked. We run into problems though, as the suggested candidate (d) is better; it violates the alignment constraint since a ϕ-phrase is found dominating a function word directly without any head (e.g., a PrWd) and is thus equally viable as candidate (a).

(25)

Func Lex_ω Func	NonRec	Align-ϕ-R	Align-PrWd-R
a. ☹ ([Func Lex_ω Func]$_\varphi$)$_I$			*
b. ([Func Lex_ω]$_\varphi$ Func)$_I$		*!	
c. ([[Func Lex_ω]$_\varphi$ Func]$_\varphi$)$_I$	*!		
d. ☞ ([Func Lex_ω]$_\varphi$ [Func]$_\varphi$)$_I$			

This suggests that one more constraint which is more highly ranked is needed to rule this possibility out: Headedness (26). What this constraint means is that every ϕ-phrase must dominate some PrWd (ω), ϕ-phrase can otherwise dominate a function word directly (which is just a foot, Σ), but only if it also dominates a PrWd. In this situation, it is only key that this constraint dominate the alignment constraint, which results in the ranking in (27).

(26) Headedness: Any C^i must dominate a C^{i-1} (except if $C^i = \sigma$). E.g., "A ϕ must dominate a ω." (Selkirk, 1995)

(27) NonRecursivity, Headedness ≫ Align-ϕ-R, Align-PrWd-R (≫ Exhaustivity)

This ranking properly rules out candidate (d) in the following tableau (28), and thus candidate (a) occurs, where the ϕ-phrase is allowed to dominate the function word as well, and that it is more important to include function words to the right than align ϕ-phrases to the right edge of PrWds.

(28)

Func Lex Func	NonRec.	Headedness	Align-ϕ-R	Align-PrWd-R
a. ☞ ([Func Lex$_\omega$ Func]$_\varphi$)$_I$				*
b. ([Func Lex$_\omega$]$_\varphi$ Func)$_I$			*!	
c. ([[Func Lex$_\omega$]$_\varphi$ Func]$_\varphi$)$_I$	*!			
d. ([Func Lex$_\omega$]$_\varphi$ [Func]$_\varphi$)$_I$		*!		*

This set of constraints also properly covers the prosodification in the sentence [berber=ʊː qɑbtɑj] 'He went to Berbera', in which a clitic attaches to a word, but the clitic does not harmonize. The reason being that the clitic is not actually part of the PrWd that it attaches to. This sentence also produces a situation in which Exhaustivity is significant and must be included; in candidate (c), the function word is dominated by an intonational phrase, which is two levels up; while ϕ-phrases are only one level up (resulting in one violation per function word). Candidate (c) is thus even worse.

(29) berber=ʊː qɑb-tɑj
berber=Foc.3pMasc go-3pM.Past
'He went to Berbera.'

(30) ([berber$_\omega$]$_\phi$ [ʊː qɑb-tɑj$_\omega$]$_\phi$)$_I$

Lex Func Lex	NonRec.	Headed.	Align-ϕ-R	Align-PrWd-R	Exhaust
a. ☞ ([Lex$_\omega$]$_\varphi$ [Func Lex$_\omega$]$_\varphi$)$_I$					*
b. ([Lex$_\omega$]$_\varphi$ [Func]$_\varphi$ [Lex$_\omega$]$_\varphi$)$_I$		*!		*	*
c. ([Lex$_\omega$]$_\varphi$ Func [Lex$_\omega$]$_\varphi$)$_I$					**!
d. ([Lex$_\omega$ Func]$_\varphi$ [Lex$_\omega$]$_\varphi$)$_I$				*!	*

The data do show some pauses, and Andrzejewski explicitly pointed out that pauses are able to delimit the spreading of the +ATR feature leftward (examples from above reproduced below in (31)). The prosodic explanation for pauses in speech is simply alignment of intonational phrases, such that each intonational

phrase is aligned to a pause. Pauses may occur at the end of the sentence, where the right edge of an intonational phrase is aligned with the right edge of an utterance; or they may occur in the middle of an utterance. This suggests the prosodification provided in (32).

(31) a. beːr-tiː mæ səː iːbsæ-tæj
garden-DEF Q DEIC buy-2PSG.PAST
'Have you bought the garden?'

b. bɛːr-tɪː # mæ səː iːbsæ-tæj
garden-DEF PAUSE Q DEIC buy-2PSG.PAST
'Have you bought the garden?'

(32) ([bɛːr-tɪː$_{\omega}$]$_{\varphi}$)$_{I}$ # ([mæ səː iːbsæ-tæj$_{\omega}$]$_{\varphi}$)$_{I}$

(33) ([bɛːr-tɪː$_{\omega}$]$_{\varphi}$ [mæ səː iːbsæ-tæj$_{\omega}$]$_{\varphi}$)$_{I}$

In order to model this, we can use a general alignment constraint, such as the following in (34), which could be ranked with the PRWD alignment already in use. In fact, for a complete ranking, several other constraints would be needed to explain the alignment of utterances to intonational phrases; and, if we were attempting to account for stress, headedness constraints (Lodovici, 2005) would also be necessary. The ranking (35) however, appears to be sufficient to explain the prosodification of the data above.

(34) ALIGN(Pause, R ; IP, R) / ALIGN-IP-R: "The left edge of every pause is aligned with the right edge of a pause."

(35) NONRECURSIVITY, EXHAUSTIVITY, HEADEDNESS ≫ ALIGN-PRWD-R, ALIGN-IP-R

3.2 Phonological Constraints

So far prosodification has been modelled based on the facts present in +ATR spreading, but the constraints haven't yet provided a model for how spreading itself works. The data and explanation so far shows that spreading occurs both directions within ϕ phrases, thus from root to suffix or suffix to root. Or in other words, if there is an +ATR vowel somewhere underlyingly in the ϕ phrase, then the whole ϕ phrase will surface as +ATR. In addition, spreading is allowed to go leftward out of ϕ phrases, and is then limited by the left edge of the intonational phrases, just as the examples in (31-33) have demonstrated with the presence and lack of a pause in speech.

One way to explain this is with alignment constraints and identity constraints. McCarthy (2003) proposes that alignment constraints can be extended to include a third category in order to count gradient violations between two edges (36). This provides the good means to model spreading, since it results in

many violations if not satisfied. To apply this to Somali, the constraint just needs the necessary variables to be inserted, as in (37).

(36) ALIGN(Cat1, Edge1; Cat2, Edge2; Cat3): ∀Cat1 if ∃Cat2, assign one violation-mark ∀Cat3 that intervenes between Edge1 of Cat1 and the nearest Edge2 of Cat2 where Cat1, Cat2 are prosodic or morphological categories, Cat3 is a prosodic category and Edge1, Edge2 ∈ {Right, Left}. (McCarthy, 2003)

(37) ALIGN(IP, L; +ATR, L; V) / ALIGN-L(IP, +ATR): For all left edges of IPs (if +ATR exists), assign one violation-mark for every vowel (V) that intervenes between the left edge of an IP and the nearest left edge of +ATR.

Thus, the constraint in (37) will mitigate leftward spreading of +ATR, by assigning one violation for every -ATR vowel between the left-most +ATR vowel and the left edge of the intonation phrase. In addition, we must formulate another constraint to ensure that spreading occurs within ϕ-phrases to the right edge of the ϕ-phrase as well, since we do see spreading rightward over items within them when the word root is +ATR.

(38) ALIGN(ϕ, R; +ATR, R; V) / ALIGN-R(ϕ, +ATR): For all right edges of ϕ-phrases (if +ATR exists), assign one violation-mark for every vowel (V) that intervenes between the right edge of a ϕ-phrase and the nearest right edge of +ATR.

The two constraints proposed are not necessarily ranked with respect to each other in any way yet, but in order to show how they work together here is an example with one of the compound words from before: [reːr-mægæːlɵ] 'townspeople' (with leftward spreading). In the tableau (39) we see that one violation is assessed in the ALIGN-L(IP, +ATR) constraint, because there is one -ATR vowel between the left-most +ATR vowel and the left edge of the IP. This also correctly applies with an individual word /dɯːʃ-ɑj/ 'she attacked' 40, because a violation is assessed when there is a -ATR vowel between the +ATR of the word root and the right edge of the ϕ phrase.

(39)

([rɛːr]$_\varphi$ [mægæːlɵ]$_\varphi$)$_I$	ALIGN-L(IP, +ATR)	ALIGN-R(ϕ, +ATR)
a. ☞ ([reːr]$_\varphi$ [mægæːlɵ]$_\varphi$)$_I$		
b. ([rɛːr]$_\varphi$ [mægæːlɵ]$_\varphi$)$_I$	*!	

(40)

[dʉːʃ-ɑj]$_{\varphi}$	Align-L(IP, +ATR)	Align-R(ϕ, +ATR)
a. ☞ [dʉːʃ-æj]$_{\varphi}$		
b. [dʉːʃ-ɑj]$_{\varphi}$		*!

The same ranking however does not yet work with a compound word in which there is an +ATR element on the left that does not spread to the right. This is because no violations are actually assessed for the alignment constraint, should the +ATR feature spread past another boundary; the constraint ranking thus far only militates for spreading of the feature up to the nearest boundary. Thus, we must use an identity constraint to assign violations to vowels which change from -ATR to +ATR, but this must be very lowly-ranked because there will certainly be violations in the event that spreading does occur. This suggests the following ranking (43), which is subsequently demonstrated in the tableau in (44):

(41) Ident(+ATR): Every +ATR vowel in the input must surface in the output.

(42) Ident(-ATR): Every -ATR vowel in the input must surface in the output.

(43) Ident(+ATR) $\gg$ Align-L(IP, +ATR), Align-R(ϕ, +ATR) $\gg$ Ident(-ATR)

(44)

([dæːn]$_{\varphi}$ [ɖɛːr]$_{\varphi}$)$_{I}$	Ident(+ATR)	Align-L(IP, +ATR)	Align-R(ϕ, +ATR)	Ident(-ATR)
a. ☞ ([dæːn]$_{\varphi}$ [ɖɛːr]$_{\varphi}$)$_{I}$				
b. ([dæːn]$_{\varphi}$ [ɖeːr]$_{\varphi}$)$_{I}$				*!
c. ([dɑ ːn]$_{\varphi}$ [ɖɛːr]$_{\varphi}$)$_{I}$	*!			
d. ([dɑ ːn]$_{\varphi}$ [ɖeːr]$_{\varphi}$)$_{I}$	*!			*

The tableau above also considers a bizarre situation in which the +ATR feature ends up on the following component in the compound, and is deleted from the initial component where it began. Once again, this ranking is demonstrated in the following tableau with an instance of leftward +ATR spreading:

(45)

	([rɛːr]$_{\varphi}$ [mægæːlə]$_{\varphi}$)$_{I}$	Ident(+ATR)	Align-L(IP, +ATR)	Align-R(ϕ, +ATR)	Ident(-ATR)
a. ☞	([reːr]$_{\varphi}$ [mægæːlə]$_{\varphi}$)$_{I}$				*
b.	([rɛːr]$_{\varphi}$ [mægæːlə]$_{\varphi}$)$_{I}$		*!		
c.	([rɛːr]$_{\varphi}$ [mɑgɑːlɔ]$_{\varphi}$)$_{I}$	*!**			
d.	([reːr]$_{\varphi}$ [mɑgɑːlɔ]$_{\varphi}$)$_{I}$	*!**			*

Now the question of +ATR spreading is fairly simple, since it only relies on the established boundaries of ϕ-phrases and intonational phrases. As such, longer examples shouldn't pose a problem. The following tableau demonstrates this with sentence (46), but note that the constraints have been abbreviated.

(46) mærkæːsʉː wæːj-æj ɔː ɪs-kɑ ɖɑqɑːq-ɑj
when.Foc.3pM fail.Past Comp Refl-From go away.Past.
'When he failed he went away.'

(47)

	([mɑrkɑːsʊː wæːj-ɑj]$_{\varphi}$ [ɔː ɪskɑ ɖɑqɑːq-ɑj]$_{\varphi}$)$_{I}$	Id(+ATR)	Align-L	Align-R	Id(-ATR)
a. ☞	([mærkæːsʉː wæːj-æj]$_{\varphi}$ [ɔː ɪskɑ ɖɑqɑːq-ɑj]$_{\varphi}$)$_{I}$				****
b.	([mɑrkɑːsʊː wæːj-æj]$_{\varphi}$ [ɔː ɪskɑ ɖɑqɑːq-ɑj]$_{\varphi}$)$_{I}$		*!		*
c.	([mɑrkɑːsʊː wɑːj-ɑj]$_{\varphi}$ [ɔː ɪskɑ ɖɑqɑːq-ɑj]$_{\varphi}$)$_{I}$	*!			

It is worth noting that there is another explanation for the feature spreading presented here, which this paper will not present a constraint-based analysis of; however it is easiest to understand the alternative after seeing the above option. This other possibility is to use a privative feature [ATR], instead of the binary system. The presence of this feature results in ATR vowels [æ e i ɵ ʉ], and the lack of it results in [ɑ ɛ ɪ ɔ ʊ]. If the latter set is not specified for ATRness underlyingly, then Identity(ATR) would perhaps be sufficient to account for why the ATR vowels always surface, but non-ATR vowels can change; the underlyingly non-ATR vowels are actually vowels without a specification.

4 Conclusion

Since the prosodification constraint set is fairly separate from the constraints governing spreading of the ATR features, and no data seems to exist yet which would prove any sort of important ranking here, it seems like the sets can be ordered in either way: Prosodification ≫ Featural, or Featural ≫ Prosodification. Although OT provides no means for derivation, it does seem tempting to rank them in the first manner suggested (48): the flexibility of the system of vowel harmony in Somali seems to suggest that spreading very far left has an amount of optionality involved. One might expect that the prosodification itself is more rigid, but the distance of spreading through leftward ϕ-phrase boundaries to the left edge of intonational phrases may sometimes not be completed if the intonational phrase is long enough... However, clear enough data for this is not yet available, but should be a good starting place for future work.

(48) NonRecursivity, Headedness ≫ Align(IP, R; ϕ, R), Align-PrWd-R(PrWd, R; ϕ, R) ≫ Exhaustivity ≫ Ident(+ATR) ≫ Align-L(IP, +ATR), Align-R(ϕ, +ATR) ≫ Ident(-ATR)

Somali Vowel Harmony is a process of spreading of [+ATR], the active feature in a binary system, which is prosodically bounded on the left by intonational phrases and on the right by ϕ-phrases. This process is easily describable using only prosodic well-formedness constraints and alignment and identity constraints.

5 Bibliography

1. Andrzejewski, B. 1955. The Problem of Vowel Representation in the Isaaq Dialect of Somali. Bulletin of the School of Oriental and African Studies. University of London (1955) vol. 17 (3) pp. 567-580
2. Edmondson, Jarold A. and John H. Esling and Jimmy G. Harris. (2004) Supraglottal cavity shape, linguistic register, and other phonetic features of Somali.
3. Krämer, Martin. 2003. Vowel harmony and correspondence theory. The Hague: Mouton de Gruyter.
4. Lecarme, Jaqueline. 1999. Focus in Somali. In the Grammar of Focus.
5. Liberman, Mark. 2002. Notes from Linguistics 202/502: Introduction to Field Linguistics. <http://www.ling.upenn.edu Accessed May 26, 2010.
6. McCarthy, John and Alan Prince. 1993. Generalized Alignment, *Yearbook of Morphology*, pp. 79-153.
7. Orwin, Martin. 1995 Colloquial Somali: a complete language course. (1995) pp. 295
8. Saeed, John Ibrahim. 1999. Somali. Amsterdam: John Benjamins.
9. Selkirk, Elisabeth. 1995. The prosodic structure of function words. In Beckman, Jill, Dickey, Laura Walsh and Urbanczyk, Suzanne (eds.) *Papers in Optimality Theory*. University of Massachusetts Occasional Papers 18. Amherst Mass: GLSA. pp. 439-469.
10. Svolacchia, M. and A. Puglielli. 1999. Somali as a polysynthetic language. L. Mereu (ed.), Boundaries of Morphology and Syntax, Amsterdam, John Benjamins, pp. 97-120.
11. Truckenbrodt, Hubert. 2007. The syntax-phonology interface. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, Cambridge: Cambridge University Press.