Title
A lexical indexation account of exceptions to hiatus resolution in Mushunguli

Author
Hout, Katherine

Publication Date
2015-06-21

San Diego Linguistic Papers **5** (2015) 30-70

A Lexical Indexation Account of Exceptions to Hiatus Resolution in Mushunguli

Katherine Hout
University of California, San Diego

Abstract

In Mushunguli (Somali Chizigula, Kizigua, ISO [xma]), vowel hiatus is typically resolved via glide formation of prevocalic high vowels (e.g. /k**u-i**v-a/ → [k**wii**va] 'to hear'), fusion of sequences of low and non-low vowels (e.g. /k**a-i**v-a/ → [k**ee**va] '(s)he heard'), or reduction of identical vowel sequences (e.g. /s**i-i**taŋg-a/ → [s**i**taaŋga] 'I called'). However, there exist a set of exceptional V-initial stems which resolve hiatus normally in glide formation and reduction contexts (e.g. /k**u-i**t-a/ → [k**wii**ta] 'to go;' /s**i-i**t-is-a/ → [s**i**tiisa] 'I went (hab.)), but exceptionally block fusion (e.g. /k**a-i**t-a/ → [k**aii**ta], *[k**ee**ta]). Two critical generalizations can be made about these stems: they all begin with high vowels, and they exceptionally block only fusion. I show that these generalizations are not accidental, but rather intrinsically linked—the exceptional stems are exceptional to fusion because fusion is the only process that materially affects the vowel of the stem, and they begin with high vowels because high vowels are the only vowels materially affected by fusion.

This paper proposes an analysis of these stems which can guarantee these generalizations. I first present an account of hiatus resolution in Mushunguli under Optimality Theory. I then propose an analysis of the exceptional stems, relying on a lexically-indexed version of IDENT(high) and a locality condition (following Pater 2010). I evaluate the predictions made by this analysis and ultimately propose a revision to the locality condition. Finally, I present and evaluate three alternative analyses: an allomorphic treatment, an abstract representational analysis, and a treatment based in Lexical Phonology. I show that although these alternatives can account for the data, they ultimately fail to capture the critical generalizations made by the lexical indexation analysis.

1 Introduction

This goal of this paper is to account for an interesting problem in the analysis of hiatus resolution in Mushunguli (Somali Chizigula; Kizigua; ISO [xma]).[1] Vowel hiatus is typically resolved in one of three ways. One is glide formation, whereby prevocalic high vowels are changed to corresponding glides, with compensatory lengthening of the second vowel (e.g. /k**u-a**sam-a/ → [k**waa**saama] 'you gaped'). Another is reduction, whereby a sequence of two identical vowels are

[1] There are a huge number of people I would like to thank for their feedback, guidance, and support in writing this paper. Thank you first to my language consultant, Mohamed Ramedhan, without whose knowledge and patience none of this would have been possible. Thank you to my primary reader Eric Bakovic, for comments on previous drafts of almost every incarnation of my work on this subject. Thank you also to my committee members Sharon Rose and Marc Garellek, who have also provided substantial guidance on this paper and other work related to the subject. Additional thanks go to David Odden, UC San Diego PhonCo and the Linguistic Field Methods Working Group. *Asante!*

reduced to a single short vowel (e.g. /k**a-a**sam-a/ → [k**a**saama] 'he/she gaped'). The third strategy is fusion, where a sequence of a low vowel followed by a non-low vowel fuses into a single long mid vowel (e.g. /k**a-i**taŋga/ → [k**ee**taaŋga] 'he/she called').[2]

Hiatus resolution is very regular in this language, and these processes are typical of Bantu languages in general (see e.g. Aoki's 1974 work on fusion in Xhosa; Clements's 1986 discussion of glide formation and compensatory lengthening in Luganda; Roberts-Kohno's 1995 work on fusion in Kikamba). However, there exists a set of stems in Mushunguli which resolve hiatus normally with regard to glide formation and reduction, but exceptionally block fusion. This is illustrated in (1) with the habitual forms of one such exceptional stem, *-it-* 'go.'

(1) *Exceptional stem behavior*

Process	*Underlying*	*Expected*	*Surface*	*Gloss*
Glide formation	/k**u-i**t-is-a/	k**wii**tiisa	k**wii**tiisa	'to go (hab.)'
Reduction	/s**i-i**t-is-a/	s**i**tiisa	s**i**tiisa	'I went (hab.)'
Fusion	/k**a-i**t-is-a/	*k**ee**tiisa	k**ai**tiisa	'(s)he went (hab.)'

While exceptional with respect to the language's general strategies of repairing word-internal vowel hiatus, these stems' behavior is strikingly systematic. As a group, two critical generalizations can be made about these stems, which in previous work (Hout 2015 (f.c.)) have been presented as in (2):

(2) *Two key generalizations*

a. The stems *all* begin with high vowels.
b. The stems are *only* exceptional to fusion.

Previous analyses of this problem have managed to capture these two generalizations satisfactorily, but have treated them as unrelated to one another (Hout 2012, Hout 2015 (f.c.)). However, further investigation indicates that these two facts are *not* unrelated, but rather intrinsically linked.

The fact that the stems begin with high vowels is not unexpected if we consider the actual results of fusion. In Mushunguli, both low+mid and low+high vowel sequences can fuse, but only in cases of low+high fusion is there any material effect on the vowel of the stem:[3]

[2] Fusion is more commonly referred to as "coalescence" in phonological literature, and is also occasionally called "contraction," especially among Bantuists. My preference for the term "fusion" is to remain consistent with previous works I have produced on this and other topics pertaining to Mushunguli.

[3] By "material effect" I am referring to changes which affect some aspect of the vowel which is contrastive in the languages. Changes in vowel quality, vowel status (i.e. changing a vowel to a consonant), and deletion can all be considered "material" changes, as all of these changes either remove a vowel entirely or alter it to a segment which appears to be another phoneme entirely. As length is not contrastive in this language, changes due to compensatory lengthening are not considered "material."

(3) *Effects of fusion*

a.	$a_1 + e_2 \rightarrow ee_{1,2}$	V_1 is raised, fronted; no change to V_2
b.	$a_1 + i_2 \rightarrow ee_{1,2}$	V_1 is raised, fronted; **V_2 is lowered**

It is not terribly surprising, then, that the *stems* which block fusion all begin with high vowels—they are the only stems which show any sort of material change to the stem-initial vowel.

The fact that the stems are only exceptional with respect to fusion is similarly intuitive, when one takes into account that fusion is the only process of the three described that actually targets and has a material effect on the stem vowel:

(4) *Targets of processes*

a.	Glide Formation	$i_1 + a_2 \rightarrow j_1aa_2$	V_1 changed to glide; V_2 unchanged
b.	Reduction	$i_1 + i_2 \rightarrow Ø_1i_2$ or $i_{1,2}$	V_1 either deleted or unchanged; V_2 unchanged
c.	Fusion	$a_1 + i_2 \rightarrow ee_{1,2}$	V_1 raised, fronted; **V_2 lowered**

In (4)a, we see that glide formation affects the first (prefix) vowel, but not the second vowel at all. In (4)b, again the first vowel is the only one that is potentially affected. But in (4)c, both the prefix *and* the stem vowel are changed. However, recall from (3) that this is only true when the stem (second) vowel is high. With these facts in mind, the generalizations from (2) do not seem to be accidents at all. Rather, they appear to be related to one another. As such, they can be better described by the revised generalizations in (5):

(5) *Key generalizations (revised)*

a. The stems *only* begin with high vowels…
…because only high vowels are materially affected by fusion.
b. The stems are exceptional to *only* fusion…
…because fusion is the only process that can materially affect a stem vowel.

Any good analysis will capture the generalizations in (2), but an excellent analysis will capture the revised generalizations in (5). The analysis below can do this, does so by using a system of lexically-indexed constraints as proposed in Pater 2010. These constraints are checked by a locality condition, which essentially serves to encode exceptionality as a property of the stem.

I provide general language background in Section 2, including the current status of the language and its speakers, data collection methods, and information about the morphology, phonemic inventory, and lengthening processes In Section 3, I will describe hiatus resolution in the language, and build an analysis in terms of Optimality Theory. In Section 4, I present the exceptional stems and their behavior in more detail, and then present the lexical indexation analysis. In this section, I also discuss interesting predictions made by lexical indexation, relate these predictions to my own analysis, and propose a potential revision to the locality condition. In Section 5, I discuss alternative analyses, with particular focus on an abstract representational solution. Finally, in Section 6, I summarize my conclusions and motivate future research.

2 Language background

2.1 Language situation

Mushunguli (Narrow Bantu, G.311), is an under-described Bantu language spoken along the lower Jubba River in Somalia. The language is related to Tanzanian Zigula, with which it is partially mutually intelligible. Modern speakers of the Mushunguli dialect are the descendants of escaped Zigula slaves who settled in the lower Jubba River Valley (Eno and Eno 2007). This community was relatively stable for over a hundred years, but was again displaced as a result of the Somali Civil War in the 1990s, during which time the Mushunguli and other Somali Bantu fled to Kenya. A large number of Somali Bantu were eventually resettled in multiple cities in the United States. Major Somali Bantu communities are located in Columbus, OH, San Diego, Boise, Buffalo, and Pittsburgh.

There are approximately 23,000 Mushunguli speakers (Lewis 2015), but given the instability faced by these people and the fact that there has been no count since 2006, this estimate may be too high. The Mushunguli people are a low-status minority (due to their former slave status) in diaspora, and interviews with members of the community indicate the beginnings of language loss. While my consultant reported that he had been teaching the language to his children, he noted that his younger children spoke it less well than his older children, and had trouble communicating with their grandmother, who spoke no English. Additionally, multilingualism with Maay and Somali is the norm within the Somali Bantu community, and Maay and Somali speakers greatly outnumber Mushunguli speakers. My consultant reported that other Somali Bantu were not interested in learning the language, and that many found it unusually challenging. Given the decreasing proficiency of younger speakers, the low status and displacement of the community, and the continuing instability in Somalia preventing any chance of return to their home, I feel it is warranted to refer to Mushunguli as an endangered language.

The data presented here are part of a corpus of over 2500 tokens that I collected during David Odden's undergraduate field methods course, and later as a member of the Ohio State University Mushunguli working group, also led by David Odden. From Spring 2011 to Winter 2012, weekly elicitation sessions were held with a native speaker, Mohamed Ramedhan. This project originally had two goals. The first was general description and documentation of the language. The second was collaboration with our consultant to develop an orthography and translate a series of folk tales from Swahili to English and Mushunguli for wider dissemination to the Mushunguli community.

Information about this language is limited, but has recently been augmented by the works produced by the OSU Mushunguli working group, including work on the semantics of locatives (Barlow 2013), noun concord in conjunctions (Williams 2012), tone (Pillion 2013), and overall language description (Odden 2013). Outside of this group, work has been done on phonetic correlates of stops (Martinez & Rosenbaum 2014) and sociophonetic variation (2013).

2.2 Phonemic inventory

Mushunguli has 22 consonantal phonemes and five vowels. The consonants are given in (6).

(6) *Mushunguli consonants*

	Labial	Labiodental	Dental	Alveolar	Post-Alveolar	Palatal	Velar	Glottal
Stop	p b			t d	tʃ	ɟ	k g	
Fricative		f v	(ð)	s (z)	ʃ			h
Nasal	m			n		ɲ	ŋ	
Trill				r				
Approx	w					j		
Lateral				l				

[ð] and [z] are in free variation in this language, hence the inclusion of parentheses. It has been reported by Martinez & Rosenbaum 2014 that [θ] and [s] have a similar pattern, but this did not appear to be the case with my consultant.

There are a number of aspects of consonant pronunciation that are not reflected in transcriptions in this paper. The primary omission is the fact that the voiced stops /b, d, g/ are typically implosive, except after non-syllabic nasals. There are also a complex set of interactions between stops and nasals. Voiceless stops following a non-syllabic nasal are generally aspirated. Word-initial nasal-stop sequences are typically pronounced as pre-nasalized stops, though Martinez & Rosenbaum 2014 have claimed that Mushunguli has a three-way distinction between word-initial pre-nasalized stops, nasal-stop sequences, and syllabic nasal+stop sequences. This potential distinction is interesting, but has not been noted or investigated in depth in my data, nor is it currently relevant to discussion of hiatus resolution. As such, my transcriptions will simply reflect pre-nasalization or syllabicity.

Mushunguli has a five vowel system, transcribed here as {a, i, u, e, o}. All non-low vowels are typically pronounced rather lax, and this is particularly true of the mid vowels, which would be narrowly transcribed as [ɛ] and [ɔ]. However, these symbols are more typically used in languages which have either more than two mid vowels (i.e. more than 3 tiers of height) or ATR distinctions. Mushunguli has neither of these qualities, so my preference is to simply make clear the three heights. Additionally, actual pronunciation of each vowel can vary due to coarticulation effects, speech rate, and speaker mood—but this is never contrastive.

Herein I will be using binary features to reference phonological vowel qualities. Other representations could be employed, but these are sufficient for the purposes of this paper. I organize the vowel space as presented in (7):

(7) *Vowel space (phonological)*

	Front ([-back])	Central (Ø)	Back ([+back])
High ([+high])	i		u
Mid ([-high, -low])	e		o
Low ([+low])		a	

The choice of [back] over [round] is arbitrary—however, the lack of specification for [back] and [round] is a deliberate choice. Roundness and backness are correlated, so roundness is predictable. I consider /a/ to be central phonetically, but phonologically placeless. However, it would not substantially alter the analysis presented here if it were represented as [-front, -back].

Tanzanian Zigula has both lexical and phrasal tone (Kenstowicz 1989, Kenstowicz & Kisseberth 1990, and Kisseberth 1992). Recent work has indicated that this may be true for

Mushunguli as well (Pillion 2013). However, tone was exceedingly difficult to hear consistently, and sometimes seemed to vary by token. Because tone does not have an effect on the processes described in this paper, it is omitted from transcriptions to avoid confusion.

2.3 Morphology

Mushunguli is similar to many other Bantu languages in that it has a fairly complex system of noun class agreement. Every noun belongs to one of thirteen noun classes, conventionally numbered 1-10, 12, 14, and 15. Classes 1-10 are grouped roughly into singular and plural pairs—non-mass lexical items belonging to the remaining three classes will take a plural counterpart from one of the plural (even-numbered) classes. Class membership is indicated by a class prefix (or lack thereof) on the noun root. Examples are given in (8).

(8) *Noun classes*

Class 1	m̩-/mw-	**m̩**gosi	'man'	*Class 2*	wa-	**wa**gosi	'men'
		mwaana	'child'			**wa**na	'children'
Class 3	m̩/mw-	**m̩**ti	'tree'	*Class 4*	mi-	**mi**ti	'trees'
		mweezi	'month'			**m**ezi	'months'
Class 5	Ø	jonda	'baboon'	*Class 6*	ma-	**ma**jonda	'baboons'
		ɟula	'frog'			**ma**ɟula	'frogs'
Class 7	tʃi-	**tʃi**fulo	'bubble'	*Class 8*	vi-	**vi**fulo	'bubbles'
		tʃisuse	'scorpion'			**vi**suse	'scorpions'
Class 9	N-/Ø	ᵐbuguni	'ostrich'	*Class 10*	N-/Ø	ᵐbuguni	'ostriches'
		suwi	'leopard'			suwi	'leopards'
Class 12	ka-	**ka**buga	'bunny'	*Class 14*	u-	**u**limi	'tongue'
		kahuɟi	'hawk sp.'			**u**tʃiza	'darkness'
Class 15	ku-	**ku**mulika	'flashing'				
		kuchema	'singing'				

The intricacies of the various forms of noun concord in this language are beyond the scope of this paper. For my purposes, however, in general any modifier or predicate of a noun must agree with that noun in class, which is done by the use of an agreement prefix. If the modifier is an adjective, generally this prefix is identical in shape to the noun class prefixes. If it is a demonstrative, possessive pronoun, or interrogative pronoun, some classes will use different prefixes, seen in (9) (using the demonstrative stem *–no* 'this'):

(9) *Demonstrative series prefixes*

Class 1	ju-	juno	*Class 2*	wa-	wano
Class 3	u-	uno	*Class 4*	i-	ino
Class 5	di-	dino	*Class 6*	ja-	jano
Class 7	tʃi-	tʃino	*Class 8*	vi-	vino

Class 9	i-	ino	*Class 10*	zi-	zino
Class 12	ka-	kano	*Class 14*	u-	uno
Class 15	ku-	kuno			

A verb must minimally contain a subject prefix, verb root, and mood suffix (either /-e/ for the subjunctive or /-a/ for the indicative; only examples with the indicative suffix will be used in this paper). Additional tense-aspect prefixes may be concatenated with the stem, as well as object prefixes (though these are only obligatory in the reflexive, or in the case of when an object is known to discourse participants but isn't explicitly mentioned). Additionally, one or more additional extension suffixes can be added to the stem before the mood suffix. Of these, only the habitual suffix *–is* will appear in examples in this paper.

In this paper, the tense/aspect combinations which will be discussed are the simple past and the present progressive. The simple past is unmarked, but takes its own set of subject prefixes for human subjects based on person and number.[4] Examples are given in (10), using the verb *hema* 'breathe.'

(10) *Simple past*

1SG	*si-*	**si**heema	1PL	*tʃi-*	**tʃi**heema
2SG	*ku-*	**ku**heema	2PL	*mu/m̩-*	**m̩**heema
3SG	*ka-*	**ka**heema	3PL	*wa-*	**wa**heema

The present progressive is marked by a different set of singular subject prefixes and the addition of a progressive prefix *a-* after the subject prefix. This triggers hiatus resolution with the vowel of the prefix; it also can trigger hiatus resolution with V-initial verb stems.[5] This is illustrated in (11), again with *hema*.

(11) *Present progressive*

1SG	*ni-*	**n**aheema	1PL	*tʃi-*	**tʃ**aheema
2SG	*u-*	**w**aaheema	2PL	*mu-*	**mw**aaheema
3SG	*a-*	**a**heema	3PL	*wa-*	**wa**heema

[4] To my knowledge, outside of class 1 and 2 (which are equivalent to 3rd sg and 3rd pl agreement markers, as classes 1 and 2 contain human lexical items only), there are no variant subject prefixes for the other classes. Rather, they simply have the same shape as the demonstrative series prefixes—though whether this means they are actually the same prefix is questionable.

[5] Because hiatus resolution will always occur in these cases, it would not be unreasonable to propose that none of these prefixes have the underlying shape given in (11), but rather only shapes of the bolded forms. Assuming this would not substantially change anything in the analysis, other than that some assumed triplet or quadruplet sequences have one less underlying vowel. It is worth noting, however, that these prefixes are basically identical in shape to the future subjunctive subject paradigm, which is marked with a prefix *na-* before the subject prefix. In these cases, before C-initial stems the shapes are identical to the italicized forms, with the exception of the 2nd singular prefix, which fuses with the vowel of the future prefix (na-u → noo). It is this consistency of paradigm that I am trying to capture by assuming the underlying forms given in (11).

2.4 Lengthening processes

Vowel length is not phonemic in Mushunguli. Surface long vowels do occur under two circumstances, however. The most common cause of a surface long vowel is a general process of penultimate lengthening, which is most likely a method of marking phrase boundaries, similar to Chichewa (Downing & Pompino-Marshall 2011). In inflected verb phrases, the penultimate syllable is significantly longer than other syllables in the word (Hout 2014). This is also generally true for isolated noun *phrases*, but is not necessarily always true for nouns elicited in isolation. In very long utterances or rapid speech the syllables are impressionistically more even in length, though this has not yet been carefully investigated. However, the majority of examples in this paper are inflected verb phrases, and as such penultimate lengthening does apply.

The second situation in which surface long vowels can be found is as a result of compensatory lengthening due to some hiatus resolution processes. Compensatory lengthening only occurs following glide formation or fusion—reduction does not cause it. That is, it only occurs in situations where the vowel is materially affected. Additionally, compensatory lengthening will not occur if it would affect the vowel in the final syllable of an utterance. For example, when the phrase [simba iijo] 'that lion,' is spoken in isolation, compensatory lengthening does not occur on the final syllable, despite the fact that the second word is derived from /i-i-o/ (a complex concatenation of the definite pre-prefix, the class 9 demonstrative series prefix, and the demonstrative root). However, it should be noted that my data primarily consists of short phrases and not sentences, so whether or not the *o* in this phrase would be long when embedded into a sentence is not something I have a clear answer to. Assuming that compensatory lengthening is tied to resolution of vowel hiatus across morpheme boundaries and *not* the syllable's position in the phrase or its focus, I would assume that it would apply in such a case.

The result of either penultimate or compensatory lengthening is roughly the same with respect to length—one process does not seem to produce a substantially longer vowel than the other. In situations where penultimate lengthening *and* compensatory lengthening could in principle both apply, the result is no longer than in situations where only one was applicable. I always transcribe length with a double vowel, regardless of its source.

3 Hiatus resolution

Vowel hiatus arising from word formation is generally disallowed in Mushunguli, and is resolved in one of three ways. The first is glide formation, whereby prevocalic high vowels are changed to a corresponding glides. The second is fusion, whereby a low vowel and a non-low vowel coalesce into a corresponding mid vowel. The third is reduction, by which sequences of identical vowels are reduced to a single vowel. Unlike the first two processes, reduction does *not* result in compensatory lengthening—the resulting vowel will be short unless it is in the penultimate syllable.

The general constraint driving hiatus resolution that I will adopt is the markedness constraint NOHIATUS (12).

(12) **NOHIATUS** (*V.V): assign a violation for every instance of two immediately adjacent vowels

This constraint functions differently than the commonly adopted ONSET, which simply penalizes onsetless syllables. I adopt NOHIATUS over ONSET because onsetless syllables do not exactly seem to be penalized in Mushunguli—at the very least, word-initial ONSET violations are never repaired (e.g. /i-N-bwa # i-ɟ-a/ → [imbwa iiɟa] ‘the dogs ate’). Sequences of adjacent vowels across morpheme boundaries are nearly always subject to repair, leaving me to assume that this is the actual cause of the processes to be discussed. Thus, I consider NOHIATUS to be a more “accurate” choice of constraint (see Orie & Pulleyblank 1998 for additional discussion of this constraint).

NOHIATUS must be highly ranked, given the regular pattern of hiatus resolution seen in the language. Each of these processes will now be explored in turn.

3.1 Glide formation

When an underlying high vowel is followed by a non-identical vowel, it becomes a corresponding glide on the surface. This process is regular and occurs in most /V+V/ and /CV+V/ contexts. Examples of glide formation are given in (13).

(13) *Examples of glide formation*

a.	/**u-e**di/	**wee**di	‘good (cl. 3)’
b.	/**u-a**-hem-a/	**waa**heema	‘you are breathing’
c.	/**u-o**ger-a/	**woo**geera	‘it (cl 3) swam’
d.	/**i-o**se/	**joo**se	‘all/the whole (cl 9)’
e.	/**i-e**reker-a/	**jee**rekeera	‘it (cl 9) floated’
f.	/**i-u**mbal-a/	**juu**mbaala	‘it (cl 9) is piled up’
g.	/k**u-a**sam-a/	k**waa**saama	‘to gape’
j.	/k**u-i**gut-a/	k**wii**guuta	‘to satiate’
h.	/m**u-e**rek-a/	m**wee**reeka	‘you pl are born’
i.	/m**u-i**taŋga/	m**wii**taaŋga	‘you pl called’

The change of a high vowel to a glide violates the faithfulness constraint IDENT(syllabic) (14).[6] Because I consider high vowels and corresponding glides to otherwise be featurally identical (see Rosenthall 1997 for a discussion), this is the minimal change necessary to resolve hiatus in these cases.

(14) **IDENT(syllabic)** (ID(syl)): assign a violation for every instance of a segment that is [α syllabic] in the input and [-α syllabic] in the output

*V.V must dominate IDENT(syl) to allow glides to be formed at all, as seen in the tableau in (15). Note that in this and all other tableaux in this paper, I differentiate prefix and suffix vowels by subscript P and S, respectively. This is done in order to make clear the changes undergone by prefix and stem vowels, and will become more relevant during the discussion of the exceptional stems.

[6] My choice of the feature [syllabic] here is not critical to the analysis. Changing a vowel to a glide violates some faithfulness constraint that prohibits changing a vowel to a consonant. This IDENTITY constraint is intended to penalize this change. What I *am* assuming here is that there are no featural distinctions between prevocalic high vowels and glides, aside from the fact that the former are vowels and the latter consonants—in this regard, I agree with Rosenthall 1997.

(15) *Glide formation*

/k$\mathbf{u}_{P}$-$\mathbf{i}_{S}$v-a/	*V.V	IDENT(syl)
k$\mathbf{u}_{P}$.$\mathbf{ii}_{S}$va	*!	
☞ k$\mathbf{w}_{P}$$\mathbf{ii}_{S}$va		*

Other potential “one change” candidates for hiatus resolution could exist. A vowel could be deleted (violating MAX (16)) or a consonant inserted between the vowels (violating DEP (17)).

(16) **MAX:** assign a violation for every input segment without an output correspondent

(17) **DEP:** assign a violation for every output segment without an input correspondent

Because epenthesis candidates are not chosen to resolve hiatus in these cases (or any cases), DEP must be ranked above IDENT(syllabic).

(18) *No epenthesis*

/k$\mathbf{u}_{P}$-$\mathbf{i}_{S}$v-a/	DEP	*V.V	IDENT(syl)
k$\mathbf{u}_{P}$.$\mathbf{ii}_{S}$va		*!	
☞ k$\mathbf{w}_{P}$$\mathbf{ii}_{S}$va			*
k$\mathbf{u}_{P}$.$\mathbf{Cii}_{S}$va	*!		

Because epenthesis has never been observed as a viable hiatus resolution strategy in this language, I will not be considering epenthesis candidates for the rest of the paper—by default, DEP must dominate all of the constraints violated by every winner.

Deletion candidates are also non-winners in this language, indicating that MAX also dominates IDENT(syllabic):

(19) *No deletion*

/k$\mathbf{u}_{P}$-$\mathbf{i}_{S}$v-a/	MAX	*V.V	IDENT(syl)
k$\mathbf{u}_{P}$.$\mathbf{ii}_{S}$va		*!	
☞ k$\mathbf{w}_{P}$$\mathbf{ii}_{S}$va			*
k$\mathbf{Ø}_{P}$.$\mathbf{ii}_{S}$va	*!		

Finally, there must be some constraint preventing the gliding of a stem vowel. To this end, I adopt a stem identity faithfulness constraint, STEM-IDENT(syllabic) (20).[7]

[7] My choice of constraint is somewhat arbitrary. Other constraints could mediate the choice between these candidates, depending on the exact syllabification of the candidate. If the glide is syllabified as a coda, e.g. kuj.va, then it would violate NOCODA. However, there is some evidence for potential (though

(20) **STEM-IDENT(syll) (S-IDENT(syl)):** assign a violation for every instance of a stem segment that is [α syllabic] in the input and [-α syllabic] in the output

STEM-IDENT(syllabic) can be ranked anywhere with respect to IDENT(syllabic), as gliding a prefix vowel violates only the latter constraint:

(21) *No gliding of the stem vowel*

/k**u**$_P$-**i**$_S$v-a/	IDENT(syl)	S-IDENT(syl)
☞ k**w**$_P$**ii**$_S$va	*	
k**u**$_P$**j**$_S$va	*	*!

Glide formation of the initial vowel is thus currently the preferred strategy for hiatus resolution. It incurs only a violation of IDENT(syllabic), which is the lowest-ranked constraint in the grammar.

3.2 Fusion

Fusion occurs when a low vowel precedes a non-low vowel. The result of fusion is a single, long mid vowel with the second vowel's place features (front and unround *e* if /i,e/; back and round *o* if /u,o/). While I consider all combinations of low+non-low vowels to be the same process, I will refer to occurrences of /a+i/ and /a+u/ as high fusion, and occurrences of /a+e/ and /a+o/ as mid fusion for the sake of explication. This process is illustrated by the data in (22).

(22) *Examples of fusion*

a.	/k**a-i**v-a/	k**ee**va	'he/she heard'
b.	/w**a-i**taŋg-a/	w**ee**taaŋga	'they called'
c.	/k**a-e**res-a/	k**ee**reesa	'she gave birth'
d.	/w**a-e**-jag-a/	w**ee**jaga	'they scratched themselves'
e.	/w**a-u**mbal-a/	w**oo**mbaala	'they are piled up'
f.	/k**a-o**mal-a/	k**oo**maala	'he/she dished up ugali'
g.	/tʃi-w**a-o**geð-a/	tʃiw**oo**geeða	'we frightened them'
h.	/k**a-o**ger-a/	k**oo**geera	'it (cl. 12) swam'

Fusion is a complicated process that incurs a number of violations. The raising of /a/ to a mid vowel violates IDENT(low) (23); it gaining a place feature also violates IDENT(back) (24).

extremely limited) codas in the language (e.g. /na-m-ɟ-e/ → [naamɟe] 'you pl. will eat'), though the actual syllabification of tokens like these is not clear. If the glide were syllabified as part of a complex onset in the second syllable (e.g. ku.jva), it would violate either a constraint against complex onsets or a constraint against falling-sonority consonant clusters. However, as previously mentioned, Mushunguli has been described as having complex nasal-stop onset clusters that are contrastive with pre-nasalized stops and sequences of syllabic nasal+stop (Martinez & Rosenbaum 2014). Finally, the glide could be part of a falling-sonority diphthong, e.g. ku͜jva, and thus be a violation of NODIPTHONG (or a constraint against falling-sonority diphthongs specifically). Since there do not appear to be any diphthongs in the language, this would be an equally valid choice of constraint.

(23) **IDENT(low)** (ID(low)): assign a violation to any output correspondent whose value for [low] differs from that of its input

(24) **IDENT(back)** (ID(back)): assign a violation to any output correspondent whose value for [back] differs from that of its input

In cases of high fusion, the lowering of the second vowel to a mid vowel also violates IDENT(high) (25):

(25) **IDENT(high)** (ID(high)): assign a violation to any input correspondent whose value for [high] differs from that of its input

A table showing the violations incurred by the winning candidates of mid and high fusion is given in (26):

(26) *Violations incurred by winning fusion candidates*

	Input	Output	Violations
a.	$a_1 + i_2$	$ee_{1,2}$	V_1: ID(low); ID(back)
			V_2: ID(high)
b.	$a_1 + u_2$	$oo_{1,2}$	V_1: ID(low); ID(back)
			V_2: ID(high)
c.	$a_1 + e_2$	$ee_{1,2}$	V_1: ID(low); ID(back)
			V_2: none
d.	$a_1 + o_2$	$oo_{1,2}$	V_1: ID(low); ID(back)
			V_2: none

There are three noteworthy observations in this table. The first is that the winning candidates of mid fusion violate a proper subset of the constraints violated by the winning candidates of high fusion. The second is that only in cases of high fusion does the second vowel incur any violations at all. The third is that the only constraint violated by the second vowel in cases of high fusion is IDENT(high).

Because each of these constraints is violated by the winning candidates of high fusion, all must be ranked below NOHIATUS to allow a fused winner to emerge at all. However, they do not appear to be ranked with respect to one another:

(27) *Allowing a fused candidate to win*

/k**a**$_P$-**i**$_S$v-a/	*V.V	IDENT(hi)	IDENT(low)	IDENT(back)
k**a**$_P$.**ii**$_S$va	*!			
☞ k**ee**$_{P,S}$va		*	*	*

As deletion is not a viable strategy for resolving hiatus in these cases, MAX must be ranked above all of the identity constraints. Its ranking with respect to NOHIATUS is still unclear, however.

(28) *No deletion*

/k**a**$_P$-**i**$_S$v-a/	*V.V	MAX	IDENT(hi)	IDENT(low)	IDENT(back)
k**a**$_P$.**ii**$_S$va	*!				
☞ k**ee**$_{P,S}$va			*	*	*
k**Ø**$_P$.**ii**$_S$va		*!			

In order to prevent initial high vowels from fusing, at least one of these constraints must be ranked above IDENT(syllabic). For my purposes, I choose IDENT(high), as it would be the constraint violated by said initial high vowel. This is illustrated in (29).

(29) *Prevention of fusion of initial high vowels*

/k**u**$_P$-**a**$_S$sam-a/	*V.V	IDENT(hi)	IDENT(syl)
k**u**$_P$.**a**$_S$saama	*!		
☞ k**w**$_P$**aa**$_S$saama			*
k**oo**$_{P,S}$saama		*!	

This ranking also prevents an initial low vowel from gliding to either *j* or *w*, but it does not prevent the gliding of a low vowel to a low glide. As non-high glides seem to be prohibited in this language, this could reasonably be prevented by a constraint like HIGHGLIDE (30), which penalizes segments which are [-cons, -syl, -high].

(30) **HIGHGLIDE:** assign a violation to any [-cons, -syl] segment which is [-high]

This constraint must be ranked above all of the IDENTITY constraints violated by a fused winner, as well as IDENT(syllabic). A tableau illustrating this is given in (31). Here, G is a stand-in for a low glide, rather than any glide.

(31) *No gliding of /a/*

/k**a**$_P$-**i**$_S$v-a/	HIGHGLIDE	IDENT(back)	IDENT(low)	IDENT(hi)	IDENT(syl)
☞ k**ee**$_{P,S}$va		*	*	*	
k**w**$_P$**ii**$_S$va		*	*	*	*!
k**j**$_P$**ii**$_S$va		*	*	*	*!
k**G**$_P$**ii**$_S$va	*!				*

Here, the change of /a/ to either high glide has the same violation profile as the fused winner, except that these also violates IDENT(syllabic). Changing /a/ to a low glide would be the preferred strategy, but HIGHGLIDE prevents this from occurring.

A constraint that has been left out of the ranking up until now is UNIFORMITY (32), which is an overarching constraint violated by fusing input segments (McCarthy & Prince 1995).

(32) **UNIFORMITY (**UFORM): assign a violation for every output segment with multiple correspondents in the input

The utility of this constraint has been called into question in Keer 1999. Its utility here is also questionable, as the decision of whether to fuse or not appears to be mediated by IDENTITY violations. However, if we assume that this constraint exists and is active in the grammar, it must be ranked *below* IDENT(syllabic). This is because if it is ranked above IDENT(syllabic), low vowels will prefer to change to high vowels (33) instead of fusing (34).

(33) *Incorrect ranking of UNIFORMITY*

/k**a**$_P$**-i**$_S$v-a/	IDENT(back)	IDENT(low)	UFORM	IDENT(hi)	IDENT(syl)
☹ k**ee**$_{P,S}$va	*	*	*!	*	
☞ k**w**$_P$**ii**$_S$va	*	*		*	*!

(34) *Correct ranking of UNIFORMITY*

/k**a**$_P$**-i**$_S$v-a/	IDENT(back)	IDENT(low)	IDENT(hi)	IDENT(syl)	UFORM
☞ k**ee**$_{P,S}$va	*	*	*		*
k**w**$_P$**ii**$_S$va	*	*	*	*!	

This ranking will still not cause initial high vowels to fuse, as these only violate IDENT(syllabic), illustrated in (35).

(35) *High vowels still do not fuse*

/k**u**$_P$**-a**$_S$sam-a/	IDENT(back)	IDENT(low)	IDENT(hi)	IDENT(syl)	UFORM
☞ k**w**$_P$**aa**$_S$saama				*	
k**oo**$_{P,S}$saama	*(!)	*(!)	*(!)		*

This holds true for any potential combination of /u+V/ and /i+V/, as any of these will incur at least one violation of IDENT(back), IDENT(low), or IDENT(high). The exception to this would be fusion of identical vowels.

The full ranking of constraints, as it stands, is given in (36).

(36) *Current ranking*

{MAX, DEP, NOHIATUS, HIGHGLIDE}
|
{IDENT(low), IDENT(back), IDENT(high)}
|
{IDENT(syllabic), STEM-IDENT(syllabic)}
|
UNIFORMITY

3.3 Reduction

The third form of hiatus resolution in Mushunguli is reduction, whereby identical vowel sequences are reduced to a single short vowel (unless the resultant vowel is in the penultimate syllable, in which case it will be lengthened). Examples of reduction are given in (37).

(37) *Examples of reduction*

a.	/s**i**-**i**-siɲa/	s**i**siiɲa	'I looked at it (cl. 9)'
b.	/**i**-**i**taŋg-a/	**i**taaŋga	'it (cl. 9) called'
c.	/k**u**-**u**mbal-a/	k**u**mbaala	'to be piled up'
d.	/w**a**-**a**mbiz-a/	w**a**mbiiza	'they helped'

Note from examples (37)a-c that reduction, *not* glide formation, is the preferred pattern of hiatus resolution for sequences of identical high vowels. In this regard, Mushunguli patterns like Luganda (Clements 1986), and is typical of languages with glide formation generally (see Casali 2011).

This process can be differentiated from glide formation and fusion of low vowels in that the result is not lengthened. The vowels in these tokens are not merely impressionistically short. Acoustic measurements of relevant inflected verb tokens found no significant difference between the length of antepenultimate syllable nuclei in surface forms generated from identical (high) /V+V/ sequences and surface forms generated from underlying (high) /V+C/ sequences (Hout 2014). It was also found that penultimate syllable nuclei in all cases were significantly longer than the antepenultimate syllables. This indicates both that hiatus is resolved in the /V+V/ cases (else we would expect a significant difference in length between the two cases), and that compensatory lengthening did not occur (else we would again expect the /V+V/ forms to be longer; additionally, we would likely expect the antepenultimate syllable nuclei to be comparable in length to the penultimate syllable nuclei).

Given these facts, my initial impulse is to treat reduction as deletion (i.e. a violation of MAX, thus forcing the ranking NOHIATUS >> MAX). However, as was already demonstrated in (28), MAX must outrank all constraints violated by fused candidates (including UNIFORMITY) in order to generate the correct result in cases of /a+V/ sequences. Glide formation is not an option, as IDENT(syllabic) also must outrank UNIFORMITY, as was shown in (33) and (34). Because UNIFORMITY is the lowest-ranked constraint in the grammar, and because fusion of two identical vowels violates no other IDENTITY constraints, a fused candidate will always be preferred in cases of sequences of identical vowels. This is illustrated in the tableau in (38). Note here that I include a constraint IDENTS as a stand-in for IDENT(high), IDENT(low), and IDENT(back), to illustrate their lack of violation and to save space.

(38) *Fusion of identical vowels*

/s**i**$_P$-**i**$_S$taŋg-a/	*V.V	MAX	IDENTS	IDENT(syl)	UNIFORMITY
s**i**$_P$.**i**$_S$taŋga	*!				
☞ s**i**$_{P,S}$taaŋga					*
s**Ø**$_P$**i**$_S$taaŋga		*!			
s**j**$_P$.**ii**$_S$taaŋga				*!	

Even if UNIFORMITY does not exist, the fusion of identical sequences will still be preferred, as it would incur zero violations (Keer 1999).

An open question, then, is why fusion of a low vowel with a non-identical vowel results in compensatory lengthening, but reduction of identical sequences does not. A deletion account would make this straightforward—glide formation and fusion preserve some timing unit (e.g. a mora or a V slot), while reduction removes the vowel and any associated timing units entirely.

This analysis does not, and actually *cannot* have this result. Recall from the introduction that the exceptional stems are exceptional with respect *only* to fusion—critically, they are not exceptional to reduction. If MAX is ranked below NOHIATUS (i.e. allowing deletion to occur), this will generate the correct result for exceptional stems with respect to resolution of /V_i+V_i/ sequences (39), but will force deletion in cases of /a+V/ sequences (40).

(39) *Correct resolution of exceptional /i-i/*

/s**i**$_P$-**i**$_S$t-is-a/	*V.V	MAX
s**i**$_P$.**i**$_S$tiisa	*!	
☞ s**Ø**$_P$**i**$_S$tiisa		*

(40) *Incorrect resolution of exceptional /a-i/*

/k**a**$_P$-**i**$_S$t-a/	*V.V	MAX
☹ k**a**$_P$.**ii**$_S$ta	*!	
☞ k**Ø**$_P$**ii**$_S$ta		*

This result indicates that MAX must actually be ranked above NOHIATUS, thus making deletion a non-viable solution for any hiatus resolution in this language.

This issue requires a re-thinking of how compensatory lengthening works, at the very least in this language. A possible solution, though currently unformalized, is that timing units are somehow connected to or “aware” of the featural content they are linked to. This would allow us to tie compensatory lengthening, at least in some cases, to IDENTITY violations. That is, because glide formation violates IDENT(syllabic), compensatory lengthening must occur. Similarly, because fusion violates any number of IDENTITY constraints, it too results in compensatory lengthening. However, because reduction of identical vowels *only* violates UNIFORMITY and does not otherwise materially affect either vowel in the sequence, the timing units are either allowed to fuse themselves, or be removed by some other grammatical pressure. A reimagining of compensatory lengthening in this way should not, in principle, prevent languages with deletion from having or not having compensatory lengthening (where appropriate)—this would still essentially tie it to whether only the segmental content or the segmental content and the timing unit are deleted. What it would do is allow languages like Mushunguli, which seem to prefer fusion over all other options, to be correctly analyzed.

A further interesting result of this analysis is that it essentially predicts a hierarchy of costliness with respect to types of fusion. As was already discussed, mid fusion violates a proper subset of the constraints violated by high fusion. Here, we see that reduction violates a proper subset of the constraints violated by mid fusion. Furthermore, of these three, high fusion is the only case where a violation is incurred solely by a change to V_2—in all other cases, the violations are either solely incurred by V_1 or shared between V_1 and V_2. This is illustrated in (41).

(41) *Violations incurred by different types of fusion*

Type of fusion	V_1 violations	V_2 violations
High fusion	IDENT(low), UNIFORMITY	IDENT(high), UNIFORMITY
Mid fusion	IDENT(low), UNIFORMITY	UNIFORMITY
Reduction	UNIFORMITY	UNIFORMITY

This observation will be relevant in the discussion of the behavior of the exceptional stems.

3.4 Vowel triplets

Triplet (and even quadruplet) sequences of vowels are possible in this language, and not necessarily rare. Recall from section 2.3 that the progressive marker is simply an /a-/, and that this marker occurs between the subject prefix and the stem (in cases where an object prefix is also used, it occurs between the subject and object prefixes). As such, any utterance which features a progressive form of a V-initial verb stem will by default feature a triplicate sequence, and any progressive V-initial verb with object marking on that verb will also at times result in a quadruplet sequence (but this is only true for object markers of the shape /V-/). This can create quite complicated hiatus contexts, but for the most part these are resolved fairly simply in the OT analysis presented thus far. Examples are given in (42).[8]

(42) *Triplicate sequences*

a.	/**a-a-a**sam-a/	**a**saama	'he/she is gaping'
b.	/N-simba # **i-i-o**/	simba **iijo**	'that lion'
c.	/s**i-i-a**z-a/	s**ijaa**za	'you asked them (cl. 4)'
d.	/s**i-u-i**v-a/	s**iwii**va	'I heard it (cl 3)'
e.	/m**u-a-i**taŋg-a/	m**wee**taaŋga	'you pl. are calling'
f.	/**u-a-o**mbok-a/	**woo**mbooka	'you are going far'

[8] I have very few examples of quadruplet sequences, so they will not be discussed in detail here.

A beauty of the analysis presented thus far is that it predicts the resolution strategies chosen by these triplets. Each surface result is the minimal violation pattern necessary to resolve hiatus. Cases like (42)a fuse the three identical vowels into a single vowel, merely incurring an extra UNIFORMITY violation. Cases like (42)b-d, with medial high vowels, resolve hiatus by gliding the second vowel, which only incurs a violation of IDENT(syllabic). In effect, this is no more costly than resolving hiatus with only two vowels. The violations incurred by the triplicate sequence in (42)c are illustrated in (43). Note that I do not consider deletion candidates, as they will always lose due to the ranking of MAX above NOHIATUS.

(43) *Resolution of medial high vowels*

/$\mathbf{i}_P$-$\mathbf{i}_P$-$\mathbf{a}_S$/	*V.V	IDENT(syl)	UNIFORMITY
$\mathbf{i}_P.\mathbf{i}_P.\mathbf{a}_S$	**!		
$\mathbf{j}_P\mathbf{ii}_P.\mathbf{a}_S$	*!	*	
☞ $\mathbf{i}_P\mathbf{j}_P\mathbf{aa}_S$		*	
$\mathbf{j}_{P,P}\mathbf{aa}_S$		*	*!

In (43) we see that any candidate that does not fully resolve hiatus is ruled out by NOHIATUS. Additionally, even a complicated set of processes (both reducing the identical vowel sequence and gliding the resultant vowel) is still more costly than simply gliding the medial vowel.

Cases with medial low vowels, such as (42)c-d, are more complex. The winner will both fuse and glide in order to satisfy NOHIATUS, even though this incurs a large number of violations. However, this is still less costly or otherwise preferred to incomplete resolution of hiatus or other hiatus resolution strategies. The tableau in (44) illustrates this for example (42)e.

(44) *Resolution of complex triplets*

/$\mathbf{u}_P$-$\mathbf{a}_P$-$\mathbf{o}_S$/	*V.V	ID(low)	ID(back)	ID(high)	ID(syl)	UFORM
$\mathbf{u}_P.\mathbf{a}_P.\mathbf{o}_S$	**!					
$\mathbf{w}_P\mathbf{aa}_P.\mathbf{o}_S$	*!				*	
☞ $\mathbf{w}_P\mathbf{oo}_{P,S}$		*	*		*	*
$\mathbf{u}_P\mathbf{w}_P\mathbf{oo}_S$		*	*	*!	*	

Here, again any candidates that have no fully resolved hiatus are automatically ruled out by NOHIATUS. Gliding of the medial /a/ is prohibited by IDENTITY violations (a low glide candidate would presumably be ruled out by HIGHGLIDE).

3.5 Problematic cases

There are two cases which go against the predictions of the OT grammar presented here so far. The first pattern to be discussed are situations where /Ci-V/ and /Cu-V/ sequences appear to resolve hiatus via elision rather than glide formation. The second pattern to be discussed is the failure of an underlying /e-/ prefix to resolve hiatus at all.

3.5.1 *Failure of glide formation*

As was seen in (13)g-i, prefixes with an underlying /Cu/ shape generally resolve hiatus via glide formation (e.g. /ku-asam-a/ → [kwaasaama], ex. (13)g). This is not true, however, for all situations featuring /Cu/ prefixes, nor is it ever the case for /Ci/ prefixes, as seen in (45)[9]:

(45) *Failure of expected glide formation*

a.	/s**i-o**ger-a/	s**o**geera, *s**joo**geera	'I swam'
b.	/m**i-e**zi/	m**ee**zi, *m**jee**zi	'months'
c.	/s**i-u**z-a/	s**uu**za, *s**juu**za	'I asked'
d.	/d**i-a**sam-a/	d**a**saama, *d**jaa**saama	'it (cl. 5) gaped'
e.	/k**u-o**mal-a/	k**o**maala, *k**woo**maala	'to dish up ugali'
f.	/m**u-o**ger-a/	m**o**geera, *m**woo**geera	'you pl swam'

In (45)a-d, we see that regardless of the preceding consonant, the expected *[Cj] clusters never surface.[10] Rather, the vowel appears to be elided. This is the typical pattern of hiatus resolution for all /Ci-/ prefixes in the language (including /tʃi-/, /zi-/, /ni-/, and /vi-/). In (45)e-f, we see that in situations where the second vowel in the hiatus context is /o/, /Cu-/ prefixes also prefer to elide the first vowel, again regardless of the consonant—however, this pattern only applies to resolution of underlying /Cu-o/ sequences.

Restrictions on surface CG clusters are not atypical in the world's languages, and perhaps should not be considered unexpected here (see Clements 1986 for a discussion of restriction on surface CG clusters in Luganda and other languages, as well as Mudzingwa & Kadenge 2011 for a discussion of a similar pattern in Karanga and Nambya). However, failure to glide is not the current expected result of the constraint ranking proposed so far.

Because deletion is not a possible hiatus resolution strategy, the best alternative is to assume that glides are formed in cases like these, but are not realized on the surface in these cases. Assuming this is a problem that needs to be handled by the phonology (as opposed to an issue of phonetic implementation), this would mean that there is likely a post-lexical ranking which deletes glides due to some kind of phonotactic pressure. For Mushunguli, these pressures can be represented as the constraint *COMPLEX (46), which penalizes complex onsets.

[9] The results of hiatus resolution here are transcribed as short, but these cases whether the result is long or short is impressionistically variable—e.g. there are tokens of /si-oger-a/ which sound like [sogeera] and which sound like [soogeera]. As there hasn't been a careful phonetic study done as to whether lengthening is occurring here, I am simply transcribing them with a short vowel.

[10] Some V-initial stems (usually but not always pronouns or demonstratives) taking the class 5 agreement prefix /di-/ exceptionally result in a palatalized [d] rather than simply eliding the consonant as in (45)d, e.g. /d**i-a**ŋgu/ → [**d^{j}aa**ŋgu], 'my (class 5).' Similarly, the verb *eat* is [kuuɟa] in the regular infinitive, but [kudiisa] in the habitual infinitive (presumably from /ku-di-is-a/). Unfortunately, not enough data exemplifying this pattern was collected to be able to discuss it in detail here.

(46) ***COMPLEX**: assign a violation for any onset with more than one consonant

This constraint must be ranked above MAX to allow glides to be deleted. In cases where palatalization could have in principle occurred to satisfy *COMPLEX, a constraint against palatalization, *Cj (47) mediates between the two candidates.

(47) **$*C^j$**: assign a violation to any palatalized consonant in the output which was not palatal in the input

The necessary ranking is illustrated in (48).

(48) *Post-lexical prevention of palatalization*

/**sj**a.sa.ma/	*COMPLEX	$*C^j$	MAX
sja.saa.ma	*!		
☞ **sØ**a̦.saa.ma			*
s^ja.saa.ma		*!	

The case of /Cu-o/ is somewhat more complicated. Most cases of /Cu-V/ do realize the glides, in these cases as either velarization of labial consonants or labialization of non-labial consonants. Assuming the feature [labial] can stand in for both of these phonetic implementations, this implies the existence of a constraint like MAX-LABIAL (49), which preserves the roundness feature from the vowel.

(49) **MAX-LABIAL** (MAX-LAB): assign a violation to every instance of a feature [+round] that is present in the input but not present in the output

This constraint will allow the labiovelar glides to be realized in relevant cases, seen in CR.

(50) *Labialization permitted*

/**kw**a.sa.ma/	*COMPLEX	MAX-LAB
kwa.saa.ma	*!	
☞ **k^w**a̦.saa.ma		
kØa.saa.ma		*!

For cases with round vowels, there must be a mediating constraint that penalizes labialized or velarized consonants before *o* (but presumably, not the labiovelar glide *w*). For now, I propose a constraint $*C^wO$ (51). However, the choice between realizing the glide in some form

(51) **$*C^wO$**: assign a violation for every labialized (or velarized) segment preceding a round vowel

This constraint can in principle apply to input sequences like |Cwu| from /Cu-u/, but as these sequences will always be reduced, it's questionable whether an input |Cwu| would ever reach the lexical level. Regardless, ranking this constraint above MAX-LAB will allow the correct sequences to be realized:

(52) *Correct output for stems beginning with round vowels*

/**kw**o.ma.la/	*COMPLEX	*C^{w}O	MAX-LAB
kwo.maa.la	*!		
k^{w}o.maa.la		*!	
☞ **kØ**o.maa.la			*!

3.5.2 *Mid vowels*

The previous sections have focused exclusively on vowel sequences with initial high or low vowels. This is not accidental—/V+V/ sequences with initial mid vowels are extremely rare, which makes it difficult to predict how they would surface in the regular case. These data do not arise often because all non-marginal cases of initial mid vowels require the use of the reflexive prefix /e-/, which occurs in the same position as the object prefix (and as a result will always either be the second vowel in a /V+V/ sequence or the medial vowel in a /V+V+V/ sequence). To my knowledge, there are no other agreement prefixes that end in /o/ or /e/. There is one stem, *toa* 'beat,' but whether hiatus in this case is underlying or the result of an optional process of intervocalic glide deletion is unclear (i.e. the underlying form may be /to-a/ or /tow-a/). Taking these caveats into consideration, my limited data does indicate that sequences beginning with mid vowels do not resolve hiatus. Examples are given in (53).

(53) *Lack of hiatus resolution when first vowel is mid*

	Underlying	*Surface*	
a.	/k**a-e-i**va/	[k**eeii**va]	'he/she heard him/herself'
b.	/s**i-e-a**mbiza/	[s**ea**mbiiza]	'I helped myself'
c.	/tʃ**i-e-o**gohez-a/	[tʃ**eo**goheeza]	'we frightened ourselves'
d.	/si-zi-t**o-a** /	[sizit**ooa**]	'I beat them (cl. 10)'

Casali (1996; 2011) finds that in a system like Mushunguli's, featuring three vowel heights and fusion of only low+non-low vowels, initial mid vowels should either glide or be deleted. However, the data in (53) run counter to this. If mid vowels were capable of gliding, we would likely expect the result of (53)b to be *[si**jaa**mbiza] (c.f. [si**wii**va] from /si-u-iv-a/, (42)f). Similarly, if mid vowels could be deleted, we might reasonably assume that in (53)a, the medial mid vowel would either delete and the remaining two vowels would fuse (resulting in *[keeva]), or the first two vowels would fuse to [e], which would end up being deleted (resulting in *[kiiva]).

The fact that mid vowels do not delete is already predicted by the current constraint ranking—nothing can delete in Mushunguli. A change to a mid glide would violate HIGHGLIDE, so it is also prohibited. The choice then is between changing to a high glide or fusing. As seen in (54), gliding is the currently preferred strategy.[11]

(54) *Gliding of mid vowels*

/k$\mathbf{a}_P$-$\mathbf{e}_P$-$\mathbf{i}_S$v-a/	*V.V	ID(low)	ID(back)	ID(high)	ID(syl)	UFORM
k$\mathbf{a}_P$.$\mathbf{e}_P$.$\mathbf{ii}_S$va	**!					
$\mathbf{kee}_{P,P,S}$va		*!	*	*		*
$\mathbf{kee}_{P,P}$.$\mathbf{i}_S$va	*!	*	*			*
☞ k$\mathbf{a}_P\mathbf{j}_P\mathbf{ii}_S$va				*	*	

Even if I were to propose a constraint specifically penalizing the gliding a mid vowel, fusion would become the winning candidate. Critically, leaving hiatus unresolved will never be an option.

I reiterate that these examples do not feature truly “initial” mid vowels, but rather mid vowels in medial position of a triplicate sequence, or stem-final vowels. It is thus not entirely clear whether this behavior is due to properties of the mid vowels themselves, or as a result of interaction with the other vowels in the sequence. Furthermore, given the general lack of mid vowels in this position in the data set, it is entirely possible that this prefix and this stem are simply exceptional. This is one potential solution to the problem—another would be to posit a constraint that specifically penalizes changing a mid vowel in any way. Until more data can be collected exemplifying the regular pattern of resolution of mid vowels, however, nothing definitive can be said.

3.6 Summary of observations

A summary of the hiatus resolution strategies for sequences of two vowels is provided in the following table. Sequences in italics are predictions based on other observed patterns in the language—those in parentheses are additionally considered to be unelicitable, based on the lack of mid-vowel-final prefixes in the language. Note here that I have made the assumption that as equence of two identical mid vowels will not reduce, given their resistance to hiatus resolution of any other kind.

[11] Leave aside for now the fact that the fourth candidate could never surface in its current form, as there is an independent restriction against *ji* and *wu* in the language. This will be discussed in more detail in section 5.2.

(55) *Table of observed hiatus resolution strategies*

V1↓ V2→	a	i	e	u	o
a	a	ee	ee	oo	oo
i	jaa	i	jee	juu	joo
e	e.a	e.i	*e.e*	*e.u*	e.o
u	waa	wii	wee	u	woo
o	o.a	*(o.i)*	*o.e*	*(o.u)*	*(o.o)*

4 Exceptional stems and analysis

A set of stems, all of which begin with a high vowels behave normally with respect to glide formation and deletion, but exceptionally block fusion. The imperative form of these stems, which has no prefix attached to it (i.e. is bare) always begin without any consonant, indicating that speakers likely internalize these stems as V-initial. This asymmetric behavior is illustrated by the near-minimal pair comparison table given in (56), where a regular stem *-iv-* 'hear' is compared with an exceptional stem *-it-* 'go'. Note that the reduction example also features the habitual suffix *-is*, to make compensatory versus penultimate lengthening contexts clear.

(56) *Comparison of exceptional vs. non-exceptional stems*

	Regular (*-iv-* 'hear')	**Exceptional** (*-it-* 'go')
Glide formation /ku-/ (2nd sg.)	k**u-i**v-a → k**wii**va 'you (sg.) heard'	k**u-i**t-a → k**wii**ta 'you (sg.) went'
Fusion /ka-/ (3rd sg.)	k**a-i**v-a → k**ee**va 's/he heard'	k**a-i**t-a → k**aii**ta, *k**ee**ta 's/he went'
Reduction /si-/ (1st sg.)	s**i-i**v-is-a → s**i**viisa 'I heard'	s**i-i**t-is-a → s**i**tiisa 'I went'

As seen in (56), the exceptional stems resolve hiatus normally in glide formation and reduction contexts—that is, they behave as if they are normal V-initial stems. However, the stems exceptionally fail to resolve hiatus in fusion contexts, behaving instead as if they were C-initial stems (c.f. [kuliima] 'you farmed,' [kaliima] 'he/she farmed'). This pattern is further illustrated in (57) below, where the behavior of a C-initial, regular V-initial, and exceptional V-initial stem (with the addition of the habitual suffix /-is/, again to make compensatory lengthening contexts clear) are compared.

(57) *C- and V-initial stem comparison table*

	C-initial (*-lim-* 'farm')	**Exceptional V-initial** (*-it-* + *-is* 'go (hab.)')	**Regular V-initial** *-itaŋg-* ('call')
1st sg past (/si-/)	siliima	sitiisa	sitaaŋga
Infinitive (/ku-/)	kuliima	kwiitiisa	kwiitaaŋga
3rd sg past (/ka-/)	kaliima	kaitiisa	keetaaŋga

Here, the box with a dashed line groups together the situations where regular and exceptional V-initial stems behave similarly. The solid line groups together situations where the exceptional V-initial and regular C-initial stems behave similarly.

An exhaustive list of the verb stems I have found that exhibit this exceptional behavior is provided in (58). Recall that penultimate syllables are lengthened.

(58) *Exhaustive list of exceptional verb stems*

	Imperative	*Infinitive (ku-)*	*3sg past (ka-)*	
a.	**ii**ta	k**wii**ta	k**aii**ta, *k**ee**ta	'go'
b.	**ii**ha	k**wii**ha	k**aii**ha, *k**ee**ha	'be bad/angry'
c.	**ii**mba	k**wii**mba	k**aii**mba, *k**ee**mba	'sing'
d.	**ii**ra	k**wii**ra	k**aii**ra, *k**ee**ra	'cry'
e.	**uu**sa	k**uu**sa	k**auu**sa, *k**oo**sa	'take out'
f.	**uu**mba	k**uu**mba	k**auu**mba, *k**oo**mba	'mold'
g.	**u**guula	k**u**guula	k**au**guula, *k**oo**guula	'lament'
h.	**uu**ŋga	k**uu**ŋga	k**auu**ŋga, *k**oo**ŋga	'want'
i.	**uu**ja	k**uu**ja	k**auu**ja, *k**oo**ja	'come back'
j.	**uu**ða	k**uu**ða	k**auu**ða, *k**oo**ða	'ask'
k.	**u**guuða	k**u**guuða	k**au**guuða, *k**oo**guuða	'care for a sick person'
l.	**u**juusa	k**u**juusa	k**au**juusa, *k**oo**juusa	'revive'

Additionally, a few nouns that take class 5/6 agreement (*ini* 'liver,' *izi* 'voice,' *ivu* 'ash') and one form of the word for 'two' (*idi*) behave similarly, though this is only diagnosable by the failure of fusion to apply where expected; appropriate morphophonological contexts for glide formation unfortunately do not exist. These stems are presented in (59).

(59) *Exhaustive list of exceptional non-verb stems*

	Underlying	*Expected*	*Surface*	
a.	/m**a-i**ni/	*m**ee**ni	m**aii**ni	'livers (cl. 6)'
b.	/m**a-i**zi/	*m**ee**zi	m**aii**zi	'voices (cl.6)'
c.	/m**a-i**vu/	*m**ee**vu	m**aii**vu	'ash (cl. 6)'
d.	/wa-ntu # w**a-i**di/	*want^{h}u w**ee**di	want^{h}u w**aii**di	'two people (cl. 2)'
e.	/m**a-i**no # m**a-i**di/	*m**ee**no m**ee**di	m**ee**no m**aii**di	'two teeth (cl. 6)'

These stems are phonetically indistinguishable from regular stems; that is, there is no indication that the quality of the initial high vowels of exceptional stems is significantly different from those of regular stems, nor are there any significant differences in length (Hout 2014). This means that there is no reason to assume that there is anything in the phonetic signal driving their exceptional behavior. Furthermore, while their behavior is exceptional relative to the typical pattern found in the language, it is still entirely systematic—there are no exceptions to these exceptions. Thus, it is necessary to account for these stems' behavior within the language's phonology.

4.1 Proposed analysis and predictions

The constraint ranking that has been independently justified by the regular cases cannot account for the behavior of these stems. Given that non-exceptional stems *must* fuse, my current ranking predicts that exceptional stems will as well:

(60) *Incorrect derivation*

/k**a**$_P$-**i**$_S$t-a/	*V.V	IDENT(high)	IDENT(low)	UNIFORMITY
☹ k**a**$_P$.**ii**$_S$ta	*!			
☞ k**ee**$_{P,S}$ta		*	*	*

This is a case of under-application, and there are no constraints or alternative rankings under classic Optimality Theory that will allow regular stems to fuse but exceptional ones to not fuse. It is thus necessary to adopt some representational or grammatical scheme that can differentiate the regular and exceptional stems.

To this end, I adopt lexically indexed constraints, as proposed in Pater 2010. This version of lexical indexation allows exceptional morphemes to be indexed to particular instances of markedness and faithfulness constraints. Both the exceptional morphemes and the lexically-indexed constraints are notationally distinguished in tableaux and rankings by a superscript L. These constraints can be freely ranked within the grammar, but critically, their scope is limited by a locality condition, defined in (61):

(61) **Locality condition**: lexically-indexed constraints "apply if and only if the locus of violation contains some portion of the indexed morpheme" (Pater 2010: 133).

What exactly counts as the "locus of violation" is not formalized. However, for now I will assume that this means that the violation must affect some exponent of the indexed morpheme. For example, in the toy example */bu^L-ival-a/, the prefix is the exceptional (lexically-indexed) morpheme. Resolving hiatus by deleting the vowel of the prefix would violate MAX^L (as well as MAX), as the violation affects the prefix vowel. However, deleting the vowel of the stem would only violate MAX, not MAX^L, as the stem vowel is not part of an indexed morpheme.

Accounting for the exceptional stems in Mushunguli is thus fairly simple. All that is needed is a lexically indexed copy of IDENT(high) (IDENT(high)^L) ranked above NOHIATUS. This will prevent exceptional stems from fusing, as illustrated in (62). Regular stems are not indexed and thus are not subject to this constraint, meaning that fusion in these cases will operate normally, shown in (63).

(62) *Fusion exceptionally blocked*

/k**a**$_P$-**i**$_S$t^L-a/	IDENT(high)^L	*V.V	IDENT(high)
☞ k**a**$_P$.**ii**$_S$ta		*	
k**ee**$_{P,S}$ta	*!		*

(63) *Fusion allowed in regular cases*

/k**a**$_P$-**i**$_S$v-a/	IDENT(high)^L	*V.V	IDENT(high)
k**a**$_P$.**ii**$_S$va		*!	
☞ k**ee**$_{P,S}$va			*

Furthermore, neither glide formation nor reduction will ever be blocked, as these hiatus resolution strategies do not incur violations of IDENT(high)^L. This is illustrated in the tableaux in (64):

(64) *Glide formation and reduction are not blocked*

/k**u**$_P$-**i**$_S$t^L-a/	IDENT(high)^L	*V.V	IDENT(syl)
k**u**$_P$.**ii**$_S$ta		*!	
☞ k**w**$_P$**ii**$_S$ta			*

/s**i**$_P$-**i**$_S$t^L-a/	IDENT(high)^L	*V.V	UNIFORMITY
s**i**$_P$.**ii**$_S$ta		*!	
☞ s**ii**$_{P,S}$ta			*

Aside from its relative simplicity, a further benefit of this analysis is that it captures and guarantees the generalizations from (5), restated here in (65).

(65) *Key generalizations (revised)*

a. The stems *only* begin with high vowels…
…because only high vowels are materially affected by fusion.
b. The stems are exceptional to *only* fusion…
…because fusion is the only process that can materially affect a stem vowel.

Indexing IDENT(high) does not only make reference to the height of the exceptional stems' vowels; this constraint actually prevents mid-vowel-initial stems from ever being exceptional with respect to fusion. This is because, as was discussed in (26), mid fusion does not incur any violation of IDENT(high), lexically-indexed or otherwise. This means that even if a mid-vowel-initial stem was indexed in the same way as the exceptional high-vowel stems, they would not block fusion. This is illustrated in (66).

(66) *No blocking of lexically-indexed mid-vowel stems*

/k**a**$_P$-**o**$_S$mal^L-a/	IDENT(high)^L	*V.V	IDENT(high)	IDENT(low)	UNIFORMITY
k**a**$_P$.**o**$_S$maala		*!			
☞ k**oo**$_{P,S}$maala				*	*

In a related fashion, the fact that fusion is the only process blocked is guaranteed by the locality condition defined in (61). As was first mentioned in (4), fusion is the only process that actually changes some facet of the identity of the *stem* vowel. What this means is that even if other constraints were indexed to some stems, other processes would not be blocked. For example, if IDENT(syl) were indexed to an exceptional stem, locality would prevent the blocking of glide formation:

(67) *No blocking of glide formation*

/k**u**$_P$-**i**$_S$t^L-a/	IDENT(syl)^L	*V.V	IDENT(syl)
k**u**$_P$.**ii**$_S$ta		*!	
☞ k**w**$_P$**ii**$_S$ta			*

Here, the locus of violation of IDENT(syllabic) is on the prefix, but the indexed morpheme is the stem. This difference means that glide formation will always proceed normally. Note that if the lexically indexed morpheme were the prefix, then the locus of violation of IDENT(syllabic) would be on the indexed morpheme. In this case, we would predict that glide formation would be blocked for every instance of the prefix, but not on any particular stem. This is a good prediction, given that in Swahili, most /Cu-/ shaped prefixes do resolve hiatus via glide formation, but the infinitival prefix (also shaped /ku-/) critically does not (Mohamed 2001).

A similar effect is seen with the indexation of IDENT(low). If IDENT(low) were indexed, the locus of violation would again be on the prefix low vowel, not the stem vowel:

(68) *Locality prevents blocking via IDENT(low)*

/k**a**$_P$-**o**$_S$mal^L-a/	IDENT(low)^L	*V.V	IDENT(high)	IDENT(low)	UFORM
k**a**$_P$.**o**$_S$maala		*!			
☞ k**oo**$_{P,S}$maala				*	*

This observation is true for stems beginning with high vowels as well—if IDENT(low) is indexed, then high-vowel initial stems cannot be prevented from fusing with this constraint either. While this is clearly not the correct solution for Mushunguli, it again makes the prediction that a language with exceptional low prefixes could, in principle, block fusion. In these cases we would predict that fusion would *always* be blocked by that prefix.

4.2 Predictions made by this analysis

The sum of these observations is that the fact that the exceptional stems which *only* block fusion *only* begin with high vowels is not accidental. Rather, fusion is blocked because it is the only process in the language that materially affects the stem vowels, and the only stem vowels in Mushunguli that are materially affected by fusion are high.

This in turn leads to an interesting prediction: there could be no language with hiatus resolution patterns just like the one in Mushunguli in which a set of high-vowel-initial stems could be exceptional with respect to glide formation but not fusion. This is because the prediction made by this theory is that *no* stems in Mushunguli can be exceptional to only glide formation—at least not using the constraints already proposed. Furthermore, there could be no language with an exceptional set of mid-vowel initial stems which only block fusion and nothing else. This is because mid fusion incurs a subset of the violations of high fusion, meaning that any constraints indexed to block the fusion of mid vowels must also block fusion of high vowels. These predictions are intuitively satisfying—they encode exceptionality as a property of the stem, and they restrict the types of possible exceptions to ones that actually affect said stem.

This theory of lexical indexation makes a number of other interesting predictions, due to the power of the theory to index *any* constraint and the weak formalization of the the locality condition. With respect to the locality condition, it is not entirely clear how constraints that can make reference to more than just the stem are affected.

The predictions made by allowing any constraints to be indexed—even those which are not active in the grammar—can be illustrated by examining ALIGN-L(Stem,σ) (69):

(69) **ALIGN-L(Stem,σ)** (ALIGN-L): assign a violation to any instance where the left edge of a stem is not aligned with the left edge of a syllable

This constraint always prefers hiatus to remain unresolved, as any cases where it is resolved moves the left edge of the stem to the right edge of the syllable. Candidates that have resolved hiatus are misaligned, as in (70):

(70) *Winners chosen by ALIGN-L*

Input	W	L
/ku-iv-a/	.ku.i_S.va.	.kwii_S.va.
/ka-iv-a/	.ka.i_S.va.	.kee_S.va,
/si-iv-a/	.si.i_S.va.	.si_S.va.

Obviously, this constraint is not active in Mushunguli, else hiatus might never be resolved at all. However, its existence means it could in principle be lexically indexed and then ranked above NOHIATUS. Doing so would obviously not provide us with a solution to the exceptional stems in Mushunguli, but would predict that there could be some language with a set of exceptional V-initial stems (of any shape) that would be exceptional to all forms of hiatus resolution. This is illustrated by a hypothetical indexed version of *iva*, in (71):

(71) *Indexation of ALIGN-L*

/k**a**$_P$-**i**$_S$v^L-a/	*ALIGN-L^L	*V.V	ALIGN-L
☞ k**a**$_P$.**ii**$_S$.va		*	
k**ee**$_{P,S}$.va	*!		*

/s**i**$_P$-**i**$_S$v^L-a/	*ALIGN-L^L	*V.V	ALIGN-L
☞ s**i**$_P$.**ii**$_S$va		*	
s**ii**$_{P,S}$.va	*!		*

/k**u**$_P$-**i**$_S$v^L-a/	*ALIGN-L^L	*V.V	ALIGN-L
☞ k**u**$_P$.**ii**$_S$.va		*	
s**ii**$_{P,S}$.va	*!		*

This is not a completely problematic prediction. There is at least one lexical item in Mushunguli which is exceptional to hiatus resolution generally, and that is *ona* 'see.' Though it is pronounced in all respects as a V-initial stem, glide formation and fusion never apply to it (e.g. /ku-on-a/ → [kuoona], *[koona] 'to see;' /ka-on-a/ → [kaoona], *[koona] 'he/she saw'). Multiple types of exceptionality are allowed in a single grammar, since indexation is morpheme-specific, and for every index, any constraint can in principle have a clone bearing that index (Pater 2010). If we want the phonological grammar of Mushunguli to be able to account for *ona*, then being able to index ALIGN-L is a good thing. Whether indexation of every type of ALIGNMENT constraint makes equally satisfying predictions is a topic for future research.

The predictions made by allowing the indexation of constraints that have a locus of violation on more the just the stem are a little less clear. For example, take the constraint UNIFORMITY, which has been proposed to be active in this ranking. The definition of UNIFORMITY is repeated in (72):

(72) **UNIFORMITY** (UFORM): assign a violation for every output segment with multiple correspondents in the input

This constraint makes reference to more than one input correspondent, meaning the locus of violation of the constraint could in principle be both segments undergoing fusion. However, the definition of locality does not make it clear whether a lexically-indexed constraint is violated when *some* part of the indexed morpheme is the locus of violation (i.e. the locus of violation could also include a part of a non-indexed morpheme), or when *only* part of the indexed morpheme is the locus of violation.

If we assume the former to be true, then in principle UNIFORMITY could be indexed to prevent both high and mid fusion:

(73) *Hypothetical mid-vowel blocking*

/k**a**$_P$-**i**$_S$t^L-a/	UNIFORMITY^L	*V.V	UNIFORMITY
☞ k**a**$_P$**.ii**$_S$ta		*	
k**ee**$_{P,S}$ta	*!		*

/k**a**$_P$-**o**$_S$mal^L-a/	UNIFORMITY^L	*V.V	UNIFORMITY
☞ k**a**$_P$**.o**$_S$maala		*	
k**oo**$_{P,S}$maala	*!		*

Furthermore, indexing UNIFORMITY would also block reduction in a language that, like Mushunguli, cannot delete to resolve hiatus between two identical vowels:

(74) *Blocking of reduction*

/s**i**$_P$-**i**$_S$t^L-is-a/	UNIFORMITY^L	*V.V	UNIFORMITY
☞ s**i**$_P$**.i**$_S$tiisa		*	
s**i**$_{P,S}$tiisa	*!		*

In effect, this means that indexing UNIFORMITY predicts the existence of a language very much like Mushunguli in which a set of exceptional V-initial stems (beginning with any vowel) block both fusion and reduction. This is perhaps not a bad prediction—if Mushunguli can prefer fusion as a hiatus resolution strategy, it stands to reason perhaps that some other language which also prefers fusion could have exceptions to that generalization. If such a language were to be found, then there is perhaps no issue with this prediction. However, a lingering issue is that this result weakens the generalization that exceptionality is a property of the stem, as violations of UNIFORMITY are simultaneously incurred by both the prefix and the stem vowel. This seems to undermine a key feature of lexical indexation over other solutions—the direct reference to the locus of violation.

If this is a problem we want to solve, there are some potential solutions. One comes from the limited utility of UNIFORMITY mentioned earlier. The only process in which it seems to be active at all is reduction, and it has been argued that fusion of identical elements should come for free (Keer 1999). The stance regarding compensatory lengthening taken earlier tied changes in segmental content to changes in length—removing UNIFORMITY seems like it may be able to solidify this result.

However, the removal of UNIFORMITY from CON makes the prediction that all identical sequences are resolved by essentially "free" fusion—in particular, it predicts that no language could ever reduce identical sequences any other way (as there would be no constraint for the relevant constraints violated by glide formation, epenthesis, or deletion to compete with). No typological study has been done to see if such an analysis is possible for all languages that have analyzed reduction of identical vowel sequences as deletion. If a language exists where this *must* be analyzed as deletion, then presumably UNIFORMITY (or some other constraint penalizing fusion) must also exist.

If this is the case, then a better solution to this issue would be to formalize the locality condition so that the locus of violation can *only* be on the indexed morpheme. Indexation of UNIFORMITY and other constraints making reference to multiple input or output correspondents would effectively be blocked. This solution seems rather less drastic than the removal of UNIFORMITY from CON, and as such is the one that I prefer. My proposed revised version of the locality condition is given in (75):

(75) **Locality (revised)**: lexically-indexed constraints apply if and only if the locus of violation contains ***only*** some portion of the indexed morpheme

This revision to locality will not adversely affect the analysis proposed for the exceptional stems in Mushunguli, as in these cases the locus of violation of IDENT(high) is only on the indexed stem.

5 Alternative analyses

Despite the slight issues with mid vowels and compensatory lengthening, I consider the analysis presented above to be an ideal (or close to an ideal) analysis of hiatus resolution and the exceptional stems in Mushunguli. This is because it both accounts for and *explains* the patterned exceptionality in the language. In this section, I consider three alternative analyses. The first relies on allomorph, the second uses an abstract representational solution and rule ordering, and the third relies on lexical leveling.

Each alternative can, in its own way, account for the data in the regular and exceptional cases, sometimes with slightly less difficulty than the Optimality Theory analysis. The abstract representational analysis can also account for the basic generalizations that all the stems begin with high vowels and are exceptional to only fusion. However, each alternative also brings with it its own set of flaws, odd predictions, and challenges. Most critically, *none of these analyses discussed in this section are able to link these two generalizations in the way that the lexical indexation account does*. It is for this reason that I disprefer them.

5.1 Allomorphy (Emergent Grammar)

A recent line of thinking in the phonological literature has called into question the role of a generative system at all in describing complex or exceptional phonological patterns. Instead, if a pattern can be learned by faculties outside of language entirely (e.g. frequency, Bayesian reasoning, etc.), it should be; see e.g. Archangeli & Pulleyblank 2011 for a discussion of this with respect to tone spreading and allomorphy in Kinande. Under such a line of reasoning, whether the behavior of the exceptional stems in Mushunguli should be handled by the grammar at all is questionable—rather an allomorphic treatment would be preferable.

It is not my intention here to dismiss Emergent Grammar. Appeals to allomorphy and other non-language-specific cognitive faculties may very well be warranted to explain many patterns and strange problems in human language. The previously-mentioned paper on Kinande is an excellent example of a well-motivated, satisfying allomorphic treatment. However, for the purposes of analyzing these exceptional stems in Mushunguli, such an analysis is much less satisfying.

An allomorphic analysis of the exceptional stems in Mushunguli would require every /a/-final prefix to have two allomorphs: one which can fuse, and one which cannot. C-initial stems by default would have to select for the non-fusing prefix, while V-initial stems would select for the fusing prefix. This line of reasoning would also entail that all prefixes ending in /i/ or /u/ would also have to both a gliding and non-gliding allomorph—again, all C-initial stems would choose the non-gliding allomorph, while all the V-initial stems would choose the gliding allomorph. This would also mean that in cases where two identical vowels are in sequence, the non-fusing/gliding allomorphs would have to be chosen and then reduced, or there would be yet a third allomorph which is "reducing." From this set of patterns, then, the exceptional stems would (for no easily explained reason) choose the gliding and/or reducing set of allomorphs, while also choosing the non-fusing set of allomorphs ending in /a/.

My objection to this analysis is that within it, the generalizations about the stems' shapes and behavior are completely accidental. The fact that all the stems begin with high vowels cannot be addressed at all, and the fact that fusion is blocked is simply because these stems arbitrarily select non-fusing allomorphs. No predictions could be made about what can and cannot be exceptional to fusion, and no explanation of why these stems and *only* these stems have this exceptional pattern is possible. My position is that these generalizations are *not* accidental and are extremely important, as they create a possible explanation for the exceptional pattern and make interesting, *testable* predictions. Even if these predictions are ultimately shown to be incorrect, the very act of testing them and finding them wanting would be a more useful enterprise than simply ignoring their existence.

The only prediction that a purely allomorphic account could make is that any V-initial stem in this language (or another language similar to it) could exceptionally choose any allomorph of any prefix for no apparent reason. This prediction, while interesting, is not really testable (assuming that it is acceptable to ignore obvious patterns and connections like the one seen in Mushunguli), and in any case seems to be empirically false, given that we do not see any mid or low vowel-initial stems in this language that exceptionally block only fusion. The fact that standard models of phonology have trouble dealing with exceptional cases to me indicates that the models need to be adjusted, not abandoned. I reiterate that calling on allomorphy to deal with some kinds of exceptional or surface-opaque processes is not a misguided or pointless endeavor. However, in Mushunguli these are not merely opaque, nor are they simply random exceptions: they are effectively a sub-pattern of the larger pattern in the language. The lexical indexation analysis captures this fact—an allomorphic treatment, I think, does not.

5.2 Abstract representations and ordering

The second alternative analysis relies on abstract underlying representations and (non-autosegmental) rule ordering. This is a slightly modified version of the analysis that is presented in Hout 2015 (forthcoming). This analysis is able to account for the un-linked version of the key generalizations that was given in (2).

This approach requires a completely different theoretical framework for analyzing hiatus in the language. For the sake of space, I will not walk through the data and motivation of each rule, but rather simply present the rules necessary to account for hiatus resolution generally in (76).

(76) *Rules and ordering*

1. **Fusion** $\underset{\begin{bmatrix}+\text{low}_a\\ -\text{high}_a\end{bmatrix}}{V_a} + \underset{\begin{bmatrix}-\text{low}_b\\ \alpha\text{back}_b\end{bmatrix}}{V_b} \rightarrow \underset{\begin{bmatrix}-\text{high}_a\\ -\text{low}_b\\ \alpha\text{back}_b\end{bmatrix}}{V_{a,b}}$

2. **Reduction** $V_i + V_i \rightarrow \varnothing V_i$ (i = identical place and height)

3. **Glide formation** $\underset{[+\text{high}]}{V} + V \rightarrow GV$

The use of CAD → CBD representations instead of the more commonly adopted A → B / C _ D notation, as well as the obviously non-standard subscript notation, is to allow a possible one-step formalization of fusion. Most treatments of fusion in rule notation describe it as a two-step rule, lowering the second vowel and then deleting the /a/. However, I have found no empirical evidence supporting this two-step treatment. Rather, what empirical evidence I *have* found supports the idea that fusion is in fact a single process. For example, in Kimatuumbi, fusion of prefix vowels is optional; that is, /a+i/ can surface as either [εε] or [a.i] (Odden 1996). However, it *cannot* surface as [a.ε]. This seems to support an "all-or-nothing" treatment of fusion, not a multi-step process. The subscript notation is thus used in my rule to show that the resulting vowel is the combination of two vowels, critically resulting from the donation of [-high] from the first vowel and [α back] from the second ([-low] is also donated from the second vowel, but I find this to be less important, as phonotactic restrictions against non-central low vowels could be invoked for this language).

To begin motivating an analysis of the exceptional stems under this framework, I note an interesting fact about Mushunguli that has heretofore not been necessary to discuss. There is a general lack of the GV sequences *ji*, *wu* on the surface. However, other stem-initial GV sequences do exist. Some examples are given in (77), though this is not an exhaustive list.

(77) *Attested stem-initial glide-vowel sequences*

a.	**ja**ga	'scratch (v)'	e.	**woo**	'them'
b.	**je**ga	'shoulder'	f.	**wa**wa	'wing'
c.	**jo**nda	'baboon'	g.	**wa**ŋgula	'hatch (v)'
d.	**ju**we	'stone'			

There are only two observed cases of *ji*, *wu*, and both are stem-internal: [m̩ɲa**wu]** 'cat (cl 3)' and [ᵐba**ji**di] 'hartebeest (cl. 9).' This is an extremely limited number of tokens, and while this could be an artifact of my relatively small lexicon, the lack of *ji* and *wu* has been corroborated in my personal correspondence with Michal Temkin Martinez of the Boise Language Project, which has a much larger (>10,000 word) lexicon. This implies that *ji* and *wu* may not be licit sequences in this language.

Additional evidence for this hypothesis comes from the words for 'cat' and 'hartebeest' themselves. 'Cat' appears to be onomatopoeic, and as such may be subject to fewer or a different set of phonotactic restrictions than regular words in the language (Childs 1988). Furthermore, the Kisseberth & Cassimjee lexicon of Mushunguli's sister language Tanzanian Chizigula has the listing *m-nyau* for 'cat' (though whether this represents two vowels in hiatus ([a.u]) or a diphthong ([au]) is not discernable from the text). The authors have a tendency to explicitly note when alternative pronunciations are available (e.g. listing both *ao* and *awo* in the entry for 'their'), so the lack of any explicit mention of a possible *m-nyawu* pronunciation is striking.

There is also some evidence that 'hartebeest' may be an oddly or partially assimilated loanword. The mostly likely candidate is from Somali, where the word for 'wild animal' is *bahal*. Another potential source is Arabic: 'بهيمة' (roughly transcribed as [bahema]), one translation of which is 'beast' or 'animal.' The Mushunguli people were displaced by the Arabic slave trade, and a large portion of the population is Muslim. A number of Arabic loanwords have made it into the language, including *tʃitabu* 'book' (c.f. 'كتاب' [kitaab] 'book'), *tʃai* 'tea' (c.f. 'شاي' [ʃai] 'tea'), and *bahari* 'ocean' (c.f. 'بحر' [bahala] 'sea'), though it isn't clear whether these were acquired directly from Arabic or from Arabic through Somali. Regardless of the source language, the potential loanword status is further evidenced by the fact that the Kisseberth & Cassimjee lexicon has no listing for any word like [^{m}bajidi]. Instead, the word *n-khongoni* is listed for 'hartebeest.'

Given these facts, it does not seem like a stretch to adopt the stance that *ji* and *wu* are not licit sequences in this language. As most other potential GV sequences are attested and not rare, an interesting question is why these two sequences are prohibited. One possible reason for this prohibition is that these two sequences are comprised of a sequence of two elements that are homorganic in place and height—that is, syllable attachment or syllabic status notwithstanding, *j* and *w* are featurally identical to *i* and *u*, respectively. This is why in the OT analysis, the only constraint violated by changing high vowels to corresponding glides was IDENT(syl).

One reasonable way to eliminate these illicit sequences is to propose a rule of homorganic glide deletion (abbreviated HGD). This is a process that is attested in other languages such as Kimatuumbi (Odden 1997), but in this language would essentially act as a morpheme structure constraint, deleting the glide of stem-initial G_iV_i sequences.[12] The rule is formalized in (78).

(78) **Homorganic glide deletion** $G_iV_i \rightarrow ØV_i$ i = same specification for place and height

However, this rule looks suspiciously like the reduction rule proposed in (76). The temptation to collapse these rules is very strong. This can in fact be done, as seen in (79):

(79) **Generalized reduction** $[-\text{cons}]_i\ V_i \rightarrow ØV_i$ i = same specification for place and height

It is unfortunately not possible to write this rule as the reduction of *any* two identical elements. This is because the rule must apply to both identical vowels across syllable boundaries and tautosyllabic G_iV_i sequences. However, it crucially *cannot* apply to sequences of the type $V_i.G_i$. If it did, underlying sequences such as /si-jag-a/ 'I scratched' would surface as *[sjaaga] or the more expected *[saaga] (given that Cj sequences do not surface), rather than the correct form [sijaaga].

Basic formalization of processes completed, recall from (57) that an interesting quality of the behavior of the exceptional stems is that they effectively behave as if they are V-initial in contexts where glide formation or reduction would apply, but act as if they are C-initial in fusion contexts. A reasonable hypothesis under rule ordering is that perhaps they actually *are* C-initial in their underlying form. An extension of this hypothesis would be that these stems begin with

[12] Or *all* G_iV_i sequences. This would require us to assume that the 'cat' and 'hartebeest' forms noted earlier would have to be marked as exceptions to this rule.

underlying G_iV_i sequences rather than just the high vowels. That is, the lexical items presented in (58) and (59) actually have the abstract underlying representations given in (80):

(80) *Abstract lexical representations of exceptional stems*

a.	/**ji**t/	'go'	i.	/**wu**s/	'take out'
b.	/**ji**ni/	'liver'	j.	/**wu**j/	'come back'
c.	/**ji**h/	'be bad'	k.	/**wu**mb/	'mold'
d.	/**ji**vu/	'ash'	l.	/**wu**ð/	'ask'
e.	/**ji**mb/	'sing'	m.	/**wu**gul/	'lament'
f.	/**ji**zi/	'voice'	n.	/**wu**guð/	'care for a sick person'
g.	/**ji**r/	'cry'	o.	/**wu**ŋg/	'want'
h.	/**ji**di/	'two'	p.	/**wu**jus/	'revive'

Given the rule of Generalized Reduction just proposed, we do not expect these initial glides to ever surface. However, giving these exceptional stems the abstract lexical representations in (80) allows their exceptional behavior to be accounted for through serial rule ordering. If Generalized Reduction is ordered after Fusion but before Glide Formation, then the correct surface forms for both exceptional and regular stems can be generated, as seen in the derivations in (81).

(81) *Derivations of exceptional vs. regular stems*

	Exceptional (/-jit-/)		*Regular* (/-iv-/)	
Underlying representation	/k**u-ji**t-a/	/k**a-ji**t-a/	/k**u-i**v-a/	/k**a-i**v-a/
Fusion	*n/a*	BLOCKED	*n/a*	|k**e**va|
Generalized Reduction	|k**ui**ta|	|k**ai**ta|	*n/a*	*n/a*
Glide formation	|k**wi**ta|	*n/a*	|k**wi**va|	*n/a*
Penultimate Lengthening	|k**wii**ta|	|k**aii**ta|	|k**wii**va|	|k**ee**va|
Surface representation	[k**wii**ta]	[k**aii**ta]	[k**wii**va]	[k**ee**va]

This analysis can account for the all-high and only-fusion generalizations, though in a different way than the lexical indexation analysis does. First, the fact that all of these stems "begin" with high vowels is captured by the fact that G_iV_i sequences can crucially only contain elements that are [+high] (recall that glides are [+high]). Second, these stems are only exceptional with respect to Fusion because Fusion is critically blocked by the presence of the glide in the underlying form—Generalized Reduction thus counterfeeds Fusion. Furthermore, the stems could not be exceptional with respect to Glide Formation, because Glide Formation is fed by Generalized Reduction.

The main selling point of this analysis is that it elegantly accounts for both the exceptional stems' behavior and shape, as well as the lack of surface G_iV_i sequences. It also does not suffer from the issue of representing reduction without compensatory lengthening that the lexical indexation analysis had. However, this analysis has a number of flaws that ultimately leads me to prefer the lexical indexation analysis.

The first, and weakest objection to this analysis is that it requires an absolute neutralization of abstract segments. This type of abstractness has been controversial for many years, as it requires positing an invisible set of segments that only exist for the express purpose of making the analysis work (see e.g. Kiparsky 1968). At best, this presents a learnability challenge—how do speakers know that the glides are present? This essentially unchecked power to posit abstract segments of any kind have led researchers to propose all kinds of potential segments. Some, arguably the least abstract, are segments that occasionally surface, such as those found in Nguni languages (see Sibanda 2011 for an overview) and Polish *yers* (Szpyra 1992). Abstract representations have also been proposed to include segments which diachronically existed in a language but synchronically do not (e.g. strong *i* in Inuit dialects (Compton & Dresher 2011), ʕ in Maltese (Brame 1972)). At the most abstract, empty consonant slots have been proposed, as in Kikamba (Roberts-Kohno 1995). This kind of power is seemingly unchecked, as long as one can make a good argument for the abstract segment.

The issue with this objection is that phonologists, in general, seem to agree that *some* degree of abstractness is necessary for our very enterprise to exist. The disagreement thus stems not from whether or not something can be abstract, but rather the fact that there is no particularly strong theory about what degree of abstraction is acceptable. Prior to the adoption of Optimality Theory (which is in principle entirely surface-driven) by the phonological community, the standard of acceptability seemed to be some undefined "significant" amount of empirical evidence that, combined with some notion of simplicity, critically pointed to an abstract representation. A gold standard abstract representational account seems to be the analysis of Nupe palatalization and labialization (Hyman 1970, 1973). This analysis proposed that Nupe had adopted loanwords from sister languages containing vowels that were unpronounceable in their language, but which would critically cause palatalization and labialization on those consonants. These unpronounceable segments were then neutralized to a pronounceable vowel that normally did not cause either palatalization or labialization—yet the palatalization and labialization still remained.

If Nupe is a reasonable standard of acceptability for abstractness underlying representations, then the abstractness of the analysis presented for Mushunguli is not an issue at all—it is arguably *less* abstract than Nupe. No unpronounceable segments are being proposed here, only unpronounceable *sequences*. That is, the glides posited to persist just long enough in the derivation to have the intended effect of blocking fusion otherwise exist within the language. There is no obvious complication from borrowing—the Kisseberth & Cassimjee lexicon lists nearly all of the exceptional stems found in Mushunguli, and they are critically noted to be exceptional to fusion in that language, as well. Additionally, the use of these glides for the purpose of accounting for exceptional stems further motivates an independently-required rule to account for why the sequences *ji* and *wu* do not appear on the surface in Mushunguli (never stem-initially, and extremely rarely otherwise). Finally, the abstract underlying forms of the exceptional stems provide at least the beginnings of an explanation as to why these stems behave as a group, as opposed to other, unsystematic exceptions in the language. All this taken into consideration, this analysis cannot be dismissed *solely* due to the level of abstraction invoked.

A much more convincing criticism is that the rule ordering presented is ad hoc, and not motivated by any independent consideration. There is no natural link between the two critical generalizations—under this model, while these two generalizations are accounted for, they are not in any way related. This has several implications. The first is that exceptionality is purely caused by ordering, and is not really a property of the stem. If the order of Glide Formation and Fusion were reversed, it would be possible to have a language exactly like Mushunguli where a set of high-vowel-initial stems are exceptional *only* to Glide Formation. Here, the fact that they begin with high vowels would truly be accidental.

While I consider it completely possible that these two facts are *not* related, there is some evidence that they might be. Languages like Kimatuumbi, which have an optional form of fusion, only allow prefix vowels to fuse (Odden 1996). Critically, prefixes and stem-initial vowels will never fuse. This could be argued as a pressure to preserve the identity of the stem—note that in Kimatuumbi, long and short vowels are contrastive, and compensatory lengthening does apply to the result of fusion. Any type of fusion of the stem vowel in this case would result in a material change. In Mushunguli, something similar is occurring—fusion is only blocked in exceptional cases where some aspect of the stem vowel would be materially affected. Because length is not contrastive, it is only the height and place features of the stem that can be "materially" affected. With this evidence in mind, a theory that can capture the link between the facts that the stem vowels are all high and that only fusion is blocked is preferred.

Other reasons for choosing the lexical indexation analysis over this one has to do with a weakness in the abstract representational analysis. As the ordering currently stands, resolution of hiatus in vowel triplets, particularly those featuring sequences of identical vowels is predicted to come out incorrectly. The data in (82) gives a few examples (some repeated from (42)):

(82) *Vowel triplets*

	Underlying	*Expected*	*Actual*	*Gloss*
a.	/**a-a-e**rek-a/	***a.ee**reeka	**ee**reeka	'he/she is being born'
b.	/**a-a-a**sam-a/	***a**.**a**saama	**a**saama	'he/she is gaping'
c.	/s**i-i-a**ð-a/	*s**aa**ða	s**ijaa**ða	'I asked it (cl 9)'
d.	/**i-i-o**/	***jo**	**ijo**	'that (cl 9)
e.	/**u-u-o**/	***wo**	**uwo**	'that (cl 3)'

The first example comes out incorrectly because Fusion applies, removing the context for Generalized Reduction. The second fails because Generalized Reduction presumably applies only once. The third fails because Generalized Reduction presumably reduces /i-i/ to |i|, which then glides and is ultimately deleted due to the restriction against Cj. The final two examples fail because Generalized Reduction is applied first, and then the result high vowel is glided.

These cases can be handled only if we stipulate a number of things: first, there must be some type of cyclicity/iterativity—first the innermost prefix is added and hiatus resolved, and then the outermost prefix is added and hiatus resolved. This also by default stipulates right-to-left rule application. By comparison, in the OT account, the resolution of hiatus in triplet sequences of vowels falls out naturally from the constraint ranking. Resolving an underlying sequence such as /i-i-o/ to [ijo] is the minimal change required to satisfy NOHIATUS. This is illustrated in (83):

(83) *Triplet sequences in OT*

/i_P-i_P-o_S/	*V.V	IDENT(syl)	UNIFORMITY
$i_P.i_P.o_S$	*!		
☞ $i_Pj_Po_S$			*
$i_{P,P}.o_S$	*!		*
$j_{P,P}o_S$		*!	*

The fact that the OT account handles sequences like these so simply is a benefit that cannot be ignored. It is this fact, in tandem with the explanatory and predictive power, that leads me to choose the lexical indexation analysis over the abstract representational one. However, while I disprefer the abstract representational analysis as a synchronic account of these phenomena, it may well represent a potential diachronic pathway for these stems to have ever become exceptional in the first place. This is a potential avenue for future research.

5.3 Level ordering

An alternative analysis that could potentially mitigate the undesirable effect of ad hoc rule ordering would be to reanalyze the stems' behavior in terms of Lexical Phonology. This would also allow, if desired, the abstract representations from (80) to be used in an OT grammar, provided there was a post-lexical constraint against *ji* and *wu* to eliminate these sequences. To account for both of the crucial generalizations in such an analysis, fusion would have to be a lexical process, while glide formation and reduction would be post-lexical. Fusion would thus apply normally to vowel-initial stems, but would fail to apply to the exceptional stems due to the critical presence of the glide in their underlying forms. At the post-lexical level, reduction and glide formation would apply normally in all cases.

Unfortunately, the key theoretical assumptions of Lexical Phonology doom this analysis before it ever begins. A critical (and desirable) trait of a post-lexical process is its ability to apply across word boundaries (Kiparsky 1982). However, the data in (84) indicate that glide formation and (potentially) reduction do not apply across word boundaries:

(84) *Hiatus across word boundaries*

a.	/u+lim**i**/#/**u**+u+o/	ulim**i** **uu**wo[13]	'that bow (cl. 14)'
b.	/ka+bug**a**/#/**a**+ka+o/	kabug**a** **a**ko[14]	'that bunny (cl. 12)'
c.	/vi+ŋk**o**/#/**i**+vi+o/	viŋk**o** **ii**vo	'these elbows (cl. 8)'
d.	/tʃi+tuŋgul**u**/#/**i**+chi+o/	tʃituŋgul**u** **ii**cho	'this onion (cl. 7)'

These data indicate that these processes are lexical, not post-lexical. Additionally, there is some evidence that fusion or something like fusion may be able to apply gradiently or optionally across word boundaries. In careful speech and phrases produced in isolation, fusion contexts do not resolve hiatus (e.g. /N+simb**a**/#/**i**+no / → [simb**a** **ii**no] 'this lion (cl. 9)'), but some tokens produced more quickly sound fused (e.g. [simb**e** **e**no]). However, the resultant vowel is (impressionistically) less clearly an *e* than in cases of fusion across morpheme boundaries. Given the rarity of these types of tokens, I am left to conclude that post-lexical fusion is at best a gradient or optional effect. However, given the fact that this is open to interpretation at *all* indicates that assigning fusion to the lexical level solely to allow an LP analysis to exist seems misguided. If we ignore these facts and assigning fusion to the lexical level and the other two processes to the post-lexical level, we are essentially abandoning a critical assumption made by Lexical Phonology in the first place. Furthermore, adopting this theory will still not link the two generalizations the way the lexical indexation account does. With this in mind, this analysis does not appear to be worth pursuing. Thus, I must conclude that this is not the appropriate tool for handling the challenges presented by these data.

[13] Compensatory lengthening either does not apply or is undone if it would affect the final syllable. The lengthened *uu* should be taken to indicate the result of penult lengthening on the first vowel; the second *u* has been glided.

[14] For sequences of identical vowels, it is essentially impossible to tell whether hiatus has been resolved or not here—I assume that it has not for the sake of a unified analysis.

6 Conclusion

In this paper, I have shown that there exist a set of high-vowel-initial stems which exceptionally block fusion but allow glide formation and reduction. Additionally, I have proposed and defended an analysis of these stems in terms of Optimality Theory and lexically-indexed constraints. In particular, I have focused on the indexation of a copy of IDENT(high) as the correct solution, as it both accounts for and *explains* the two generalizations I consider to be critical—all of the exceptional stems begin with high vowels, and they are only exceptional with respect to fusion.

As is typical of most in-depth investigations of any complicated phonological phenomena, my analysis raises as many questions as it solves. The most obvious question for future research is whether my proposed formalization of locality makes the correct predictions cross-linguistically. If a language was found to have a set of stems that block both reduction and fusion, but not glide formation, then this formalization is incorrect, as it would suggest that UNIFORMITY is indexed. I am unaware of the existence of any such language at this time. However, there are presumably other constraints that are violated by more than just the indexed morpheme—these need to be examined in more detail to see whether this formalization is correct.

A second, somewhat related line of future inquiry is the exact nature of what fusion is. Authors have attempted to describe its typology (Casali 1996; 2011), and have used it to support arguments for various representational schema (Schane 1987; Parkinson 1996) and phonological modules generally (Iskarous 1998). Despite this, when all types of fusion are taken into account it remains an extremely difficult process to formalize in a satisfactory way, especially when compared to (relatively) simpler phenomena such as glide formation.[15] However, it is typologically not a rare process, especially in Niger-Congo languages (Casali 1996). My research suggests that the height of the vowels is a factor, and that fusing low and high vowels is more costly than fusing mid vowels. I would argue that any full account of fusion as a process must take this into account.

A final, related question would be the nature of phonological exceptionality itself. I do not know of a human language that does not have exceptionality in some regard, but typically these are treated as noise or are done away with by theoretical devices such as abstract representations or lexical indexation, with little attention paid to the predictions made by such devices. However, the case of Mushunguli seems to suggest that at least some types of patterned exceptionality need to not only be handled, but must be *predicted* to occur. Generative phonological theories are generally built around the regular case, but it seems to me that a truly exceptional (pun intended) generative theory could predict systematic exceptions such as those seen in Mushunguli. Creating such a theory seems extremely challenging, but is a line of research that I consider worth at least attempting.

References

Aoki, P. K. 1974. An observation of vowel contraction in Xhosa. In *Studies in African Linguistics* 5: 223-241.

Archangeli, D. & D. Pulleyblank. 2011. Tonal allomorphy in Kinande. In *Capturing Phonological Shades.*

Barlew, J. 2013. Point of View in Mushunguli Locatives. In *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, ed. Olanike Ola Orie and Karen W. Sanders, 115-129. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2961.

Brame, M. K. 1972. On the abstractness of phonology: Maltese ʕ. In *Contributions to Generative Phonology*, ed. Michael K. Brame. 22-61.

[15] This is not to say that glide formation poses no interesting questions or is a “solved” problem, simply that it is more easily handled by our various representational schema.

Casali, R. F. 1996. Resolving Hiatus. Ph. D dissertation, UCLA.

Casali, R. F. 2011. Hiatus resolution. In *The Blackwell Companion to Phonology*, vol. 3, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice, 1434-1460. Malden, MA and Oxford: Wiley-Blackwell.

Childs, G. T. 1988. The phonology of Kisi ideophones. *Journal of African languages and linguistics*, 10(2), 165-190.

Clements, G. N. 1986. Compensatory lengthening and consonant gemination in Luganda. In *Studies in compensatory lengthening*, ed. by Leo Wetzel & Engin Sezer, 37-77.

Compton, R. & Dresher, B. 2011. Palatalization and "strong *i*" across Inuit dialects. *The Canadian Journal of Linguistics/La revue canadienne de linguistique*, 56.2. 203-228.

Downing, L. & Pompino-Marschall, B. 2013. The focus prosody of Chichewa and the Stress-Focus constraint: a response to Samek-Lodovici (2005). *Natural Language and Linguistic Theory*, 31(3), pp. 647-681.

Eno, O. & Eno, M. 2007. The journey back to the ancestral homeland: the return of the somali bantu (wazigwa) to modern tanzania. In Kusow, Abdi & Stephanie R. Bjork (eds.) *From mogadishu to dixon: the somali diaspora in a global context*. Trenton, NJ: Red Sea Press.

Henderson, B. 2010. Chimwiini: Endangered Status and Syntactic Distinctiveness. *The Journal of West African Languages*, 37, 1, 3-6.

Hout, K. 2012. The Vocalic Phonology of Mushunguli. B.A. Thesis, The Ohio State University. Retrievable at http://kb.osu.edu/dspace/handle/1811/53373.

Hout, K. 2014. Exceptionality in Somali Chizigula—Phonetic or Phonological? Ms., UC San Diego.

Hout, K. 2015. Exceptions to hiatus resolution in Mushunguli (Somali Chizigula). M.S. Submitted to *Africa's Endangered Languages*.

Hyman, L. 1970. How concrete is phonology? In *Language* 46. 58-76.

Hyman, L. 1973. Nupe three years later. In *Language* 49. 447-452.

Iskarous, K. 1998. Vowel dynamics and vowel phonology. Proceedings of the West Coast Conference on Formal Linguistics 17. CSLI, pp. 35-40.

Kenstowicz, M. 1989. Tone and accent in Kizigua - a Bantu language. in P.M. Bertinetto & M. Loporcaro (eds). *Certamem phonologicum: papers from the 1987 Cortona Phonology Meeting*, pp. 177-188. Torino: Rosenberg and Sellier.

Kenstowicz, M. & Kisseberth, C. W. 1990. Chizigula tonology: the word and beyond. In S. Inkelas & D. Zec(eds) *The phonology-syntax connection*, pp. 163-194. Chicago, IL: University of Chicago Press.

Kiparsky, P. 1968. How Abstract is Phonology? Reprinted as Chapter 6 of *Explanation in Phonology*, pp. 119-164. Mouton de Gruyter, 1982.

Kiparsky, P. 1982. Word-formation and the lexicon. In *Proceedings of the 1982 Mid-American linguistics conference* (Vol. 3, p. 22). Lawrence: University of Kansas.

Kisseberth, C. W. 1992. Metrical structure in Zigula tonology. in D. Gowlett (ed.), *African linguistic contributions*, pp. 227-259. Pretoria: Via Afrika.

Kisseberth, C.W. & Cassimjee, F. *Chizigula Lexicon Exemplified*. Ms., unpublished.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. Mushungulu. *Ethnologue: Languages of the World, Seventeenth edition.* Dallas, Texas: SIL International. Online version: http://www.ethnologue.com

Martinez, M. and V. Rosenbaum. "Acoustic and Aerodynamic Data of Somali Chizigula Stops." ACAL 45. Lawrence, KS. 18 April, 2014.

McCarthy, J. and A. Prince. 1995. Faithfulness and Reduplicative Identity. In UMOP 18: Papers in Optimality Theory, ed. J. Beckman, S. Urbanczyk and L. Walsh, 249-384. Amherst, Mass.: GLSA. ROA-60.

Mohamed, M. A. 2001. *Modern Swahili Grammar*. East African Publishers.

Mudzingwa, C. & Kadenge, M. 2011. Comparing hiatus resolution in Karanga and Nambya: an Optimality Theory Account. In *Nordic Journal of African Studies* 20(3). pp. 203-240.

Odden, D. 1996. *The Phonology and Morphology of Kimatuumbi*. Oxford: Oxford University Press.

Odden, D. 2011, May 5. *The Mushunguli – Chizigua language of Somalia.* Retrievable at http://www.ling.ohio-state.edu/~odden/mushunguli/

Orie, Olanike Ola & Douglas Pulleyblank. 1998. *Vowel elision is not always onset-driven*. Unpublished ms., Tulane University & University of British Columbia. Retrievable at http://roa.rutgers.edu/files/290-1298/roa-290-orie-1.pdf

Parkinson, F. B. 1996. The representation of vowel height in phonology. PhD dissertation, The Ohio State University.

Pater, J. 2010. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Steve Parker, (ed.) *Phonological Argumentation: Essays on Evidence and Motivation.* London: Equinox. 123-154.

Pillion, E. 2013. An Autosegmental Analysis of Verbal Tone in Mushunguli. Honors Thesis, The Ohio State University. Retrievable at: http://kb.osu.edu/dspace/handle/1811/54798

Roberts-Kohno, R. R. 1995. Vowel coalescence and hiatus in Kikamba. In *Theoretical Approaches to African Linguistics*, ed. by Akinbiyi Akinlabi, 313-327. Trenton, New Jersey: Africa World Press.

Rosenthall, S. 1997. The distribution of prevocalic vowels. In *Natural Language & Linguistic Theory* 15.1. 139-180.

Schane, S. 1985. The Fundamentals of Particle Phonology. In Phonology Yearbook , vol. I. C. Euren and J. Anderson, eds. pp. 129-155. Cambridge University Press.

Sibanda, G. 2011. Ghost segments in Nguni. In *Select Proceedings of the 40th Annual Conference on African Linguistics*, eds. Eyamba G. Bokamba et al. 130-145. Somerville, MA: Cascadilla Proceedings Project.

Szpyra, J. 1992. Ghost segments in nonlinear phonology: Polish yers. In *Language*, 68.2. 277-312.

Tse, H. 2013. Methodological considerations in the study of sociophonetic variation in an under-documented minority language: Somali Bantu Kizigua as a case study. *Proceedings of Journées d'Études Toulousaines*, University of Toulouse, 129-139.

Williams, S. T. 2012. Gender Conflict Resolution in Mushunguli. B.A. Thesis, The Ohio State University. Retrievable at: http://kb.osu.edu/dspace/handle/1811/52072

Katherine Hout
University of California, San Diego
khout@ucsd.edu