The Diachronic Emergence of Retroflexion in Somali Bantu Kizigua: Internal Motivation or Contact-Induced Change?

Holman Tse

1. Introduction

Somali Bantu Kizigua (SBK) is an under-documented and possibly endangered variety of the Tanzanian language Zigula (or Zigua), classified by Bantu linguists as G.31 (Guthrie 1967). This dialect has also been referred to by its Somali name, Mushungulu (G.311)[1]. It is a language that has recently come to the attention of linguists through Somali Bantu[2] refugee communities that have arrived in the US in the past decade (see Odden n.d.; Barlew 2013; Martinez 2013; Tse 2013). In one of the few available pre-Somali Civil War studies, Crevatin (1993) described "Cizigula" as "una lingua ritrovata" (a rediscovered language) and as one that has generally been well maintained in spite of an earlier history of migration. As will be argued in this paper, the migration from Tanzania to Somalia in the 19th Century, which resulted in contact with speakers of Cushitic and other Bantu languages, proves to be crucial in understanding a typologically unusual set of sound changes. The changes in question have all led to the emergence of retroflex stops, which are unattested in Tanzanian Zigula (TZ). This paper will focus on whether retroflexion developed due to internal or external motivation.

SBK has three retroflex stops, which are all articulated with varying degrees of tongue curling. These include a retroflex implosive, /ɖ/, and two retroflex NC (homorganic nasal + consonant) complexes, /ɳʈ/ and /ɳɖ/, contrasting in voicing. Phonetically, retroflex NC is accompanied by a trill-like aspiration and is prenasalized in some contexts. As in the case of many Bantu languages, the phonological status of NC in SBK is ambiguous. While in some contexts it behaves like a single consonant, in others it behaves like two distinct consonants. For the purpose of this paper, NC will be treated as a single phonemic unit, but transcribed as nasal + consonant as in /ɳʈ/ and /ɳɖ/ following common Bantuist transcription conventions[3]. The two retroflex NC stops also contrast with two alveolar NC stops. All of these coronal stop contrasts are illustrated in Table 1.

What makes the SBK case remarkable is that there is evidence for both internal and contact-induced change. On the one hand, the highly regular pattern found in the diachronic data makes one suspect internal motivation. On the other hand, it is hard to ignore the history of migration and the are-

[1] The most recent classification of Bantu languages (Maho 2009) lists Mushungulu (or 'Mushunguli') in Somalia as G.311 and hence a separate but closely related language. Ethnologue (Lewis et al. 2013), however, leaves the question of mutual intelligibility an open one. For the purpose of this paper, the varieties spoken in Tanzania and in Somalia will be considered different dialects of the same language.
[2] See Besteman (2012) for a discussion of "Somali Bantu" as an ethnonym. Nurse (2010) uses it in a linguistic sense to refer to Bantu languages spoken in present-day Somalia.
[3] Although NC is phonetically prenasalized in some contexts, the full range of its phonological behavior seems better captured using this notation rather than with superscripts as in /ⁿʈ/ and /ⁿɖ/.

al linguistics research showing retroflexion to be an areal feature in Somalia, but not in Northeastern Tanzania. This paper will argue that if we consider the social history of the Zigua people and the specific groups with which they interacted after arrival in Somalia, we can see that retroflexion is more likely due to contact and that what is most surprising is that it was more likely due to contact with other Bantu languages historically spoken in Southern Somalia than with Cushitic languages, which are otherwise known for having retroflex stops.

Table 1: Examples Illustrating Alveolar/Retroflex Contrast

Sound	TZ (Kisbey 1906)	TZ (Mochiwa 2008)	SBK	Sound Change	Gloss
t	ma**t**unda	ma**t**unda	matun**ɖ**a	--	'fruits' (pl)
ɗ	ma**d**udu	ma**d**udu	ma**ɗ**uɗu	d > ɗ	'bugs' (pl)
ɳʈ	wa**nt**u	wa**nt**u	wa**ɳʈ**u	nt > ɳʈ	'people' (pl)
nt	--	--	hara**nt**i	--	'courtyard'
ɳɖ	-nku**nd**u	-enku**nd**u	-hu**ɳɖ**u	nd > ɳɖ	'red'
nd	--	--	**nd**oni	--	'boat'

In developing an account based on contact-induced change, this paper adopts the framework established by Thomason and Kaufman (1988) who argue that external factors are always primary. This is the reverse of the traditional historical linguistics approach, which leans more strongly in favor of seeking internal over external causes for change. As Thomason and Kaufman have said, "the history of a language is a function of its speakers, and not an independent phenomenon that can be thoroughly studied without reference to the social context in which it is embedded" (1988:4).

After presenting the data in fuller detail in the next section, this paper will review the typological literature on retroflexion. This will be followed by a discussion of the historical contact situation in southern Somalia in Section 4. Section 5 will discuss the languages for which we have evidence of contact and argue that the greater similarity in phonological structure in Bantu languages makes these languages a more likely source. Section 6 develops an analysis of how each retroflex stop may have developed by synthesizing material from the previous sections while Section 7 concludes this paper. Since retroflexion is not very common, an analysis of its emergence in SBK has broader implications for the development of knowledge about what kind of phonological systems are possible and about how they can change and develop over time. Furthermore, the unique history of migration involved provides fertile ground for researchers interested in the various outcomes of language contact.

2. The Data

The data for this paper comes primarily from a missionary-produced dictionary of TZ containing approximately 3,500 words (Kisbey 1906) and from consultant work with present-day speakers from a Somali Bantu refugee community in the United States. These speakers include a 21-year old female native speaker who worked with the author as part of a Field Methods course at the University of Pittsburgh that took place from January to April 2012. Three additional speakers were later recruited as part of a broader survey of speech patterns. All three were male between the ages of 22 and 30. All four speakers produced retroflex sounds in the same environments. Other sources of data include a more recent dictionary of TZ (Mochiwa 2008) and an online lexicon of SBK (Odden n.d.).

The complete consonant inventory of SBK is shown in Table 2 below and was compiled based on a lexicon of approximately 700 words. This includes all 220 basic vocabulary items listed in Samarin (1967). By identifying corresponding forms in TZ as found in (Kisbey 1906), several sound changes have been identified as indicated in Table 2. From all of this data, we can see that the three retroflex stops in SBK emerged as a result of three sound changes:

(i) [-nasal] plosives > implosives, ɗ > ɗ / _ [V, +back]
(ii) nt > ɳʈ
(iii) nd > ɳɖ

The first sound change is one that actually involves an intermediate step. All plosives in TZ have become implosives in SBK. Once this change occurred, retroflex [ɖ] developed from alveolar [ɗ] when preceding back vowels (/u, o, a/) while alveolar [ɗ] is produced preceding front vowels (/i, e/). The difference between [ɖ] and [ɗ] is, thus, allophonic and not contrastive and holds even for loan words such as [ɗunia]. Words illustrating this pattern are shown in Tables 3 and 4.

Table 2: SBK Consonant Inventory with Changes from TZ in Bold

		Lab.	Lab-dent.	Coronal		Palatal	Velar	Uvular / Glottal[4]
				Alveolar	**Retroflex**			
[-voi] stops	[-nasal]	p		t		**k > c / _ i**	k	
	[+nasal]	mpʰ		**(nt)**	**ntʰ > ɳʈ**			**nkʰ > nq**
[+voi] stops	[-nasal]	**b > ɓ**		**d > ɗ / _ [V, -bk]**	**d > ɖ / _ [V, +bk]**	**dj, gi, ge > ʄ**	**g > ɠ**	
	[+nasal]	mb		**(nd)**	**nd > ɳɖ**		ŋg	
Fricatives			f v	s z		ʃ		h
Approximants		w		l		j		
Nasals		m		n		ɲ	ŋ	
Trills				**l > r**				

Given the lack of a contrast between non-NC stops (represented as [-nasal]) and implosives both historically and in the present, one may wonder if voiced stops in TZ were actually implosives but not noted as such. All available sources on TZ show the presence of plosives to the exclusion of implosives. This includes Kisbey (1906), Brenzinger (1987), Kenstowicz (1988), Guthrie (1967), which otherwise noted a distinction between plosives and implosives in other Bantu languages but described TZ as having plosives, and Mochiwa (2008), which adopted IPA symbols. In contrast, sources on SBK are unanimous in describing the voiced stops as implosives (Crevatin 1993; Odden n.d.). Thus, the evidence shows that at some point in time, inherited plosives in TZ became implosives in SBK.

Also, since there is no phonological contrast between [ɗ] and [ɖ], one could argue that retroflexion is simply a phonetic feature with no independent phonological status. When placed in the context of the larger coronal stop inventory, however, we can see that the other two retroflex stops in SBK do contrast with alveolar counterparts. The consonant inventory as a whole, thus, motivates an analysis of the language as one containing three retroflex stops. So for this reason and the fact that the retroflex [ɖ] occurs in a greater number of environments (preceding 3 back vowels as opposed to 2 front vowels)[5], /ɖ/ is treated as basic rather than /ɗ/.

The other two sound changes resulting in retroflexion both involve NC. In all cases of inherited vocabulary in which we find a /t/ or /d/ in TZ following /n/, we find retroflex stops in SBK. Some examples are illustrated in Table 5. Exceptions to post-nasal retroflexion appear to be limited to non-inherited vocabulary as illustrated in Table 6. Thus, it appears that loan words can contain alveolar /nt/ and /nd/ while cases of inherited /nt/ and /nd/ all became /ɳʈ/ and /ɳɖ/. The alveolar /nt/ and /nd/ have, thus, reemerged in the language. The overall result of retraction of inherited NC sequences and the subsequent borrowing of vocabulary with alveolar NC sequences is a consonant inventory that includes an alveolar/retroflex contrast that is absent in TZ. An example of a near minimal pair would be [ɓaɳʈi] vs. [haranti].

These sound changes show NC behaving as a single phonological unit. When looking at the actual phonetic articulation, however, it is not clear that this is always the case. For instance, in word-initial position, /ɳʈ/ and /ɳɖ/ are produced as prenasalized stops accompanied by a trill-like aspiration and hence as [ⁿʈʳ] and [ⁿɖʳ]. In word-medial position, however, it is uncertain whether or not they are also prenasalized or articulated as separate segments with /n/ syllabified as part of the preceding syllable. A

---

[4] For some speakers in some contexts, there is an additional change: nq > h. See Tse (2013) for further details.

[5] The possible environments in which implosives can occur are limited due to a 5-vowel inventory and a (C)V syllable structure (ignoring the complications raised by the analysis of NC).

moraic interpretation also seems possible. There is similar behavior for all NC stops in SBK. The full range of synchronic behavior displayed by NC is quite complex and is worth further investigation.[6] For the purpose of this paper, however, the crucial fact is that NC behaves diachronically as a single unit. A diachronic change in its phonetic realization has had a phonological impact by making possible an alveolar/retroflex contrast absent in TZ.

Having presented the data on retroflexion, we now turn to the question of how it developed. The next section addresses the possibility of internal motivation by reviewing the typological literature on the emergence of retroflexion.

Table 3: The Coronal Implosive Preceding Back Vowels

TZ (Kisbey 1906)	TZ (Mochiwa 2008)	SBK	Gloss
ku**d**anta	ku**d**anta	ku**ɗ**aɳʈo[7]	'to lie'
dole	**d**ole	**ɗ**ole	'finger'
-**d**o**d**o	-**d**o**d**o	-**ɗ**o**ɗ**o	'small'
du**d**u	**d**u**d**u	**ɗ**u**ɗ**u	'bug'
--	--	mi**ɗ**u**ʄ**ali	'woods' (origin unknown)
--	--	**ɗ**unia	'world' (from Standard Swahili)
ku**d**umula	--	ku**ɗ**umula	'to cut'

Table 4: The Coronal Implosive Preceding Front Vowels

TZ (Kisbey 1906)	TZ (Mochiwa 2008)	SBK	Gloss
-e**d**i	-e**d**i	-e**ɗ**i	'good'
dihi	--	**ɗ**ihi	'which'
kigu**d**i	cigu**d**i	ci**ʄ**u**ɗ**i	'hip'
kai**d**i	--	kai**ɗ**i	'two'
m̩**d**ege[8]	**d**ege	**ɗ**e**ʄ**e	'bird'

Table 5: nt > ɳʈ and nd > ɳɖ in Inherited Vocabulary

Late 19th Century TZ	SBK	Gloss
m̩**nt**u	m̩**ɳʈ**u	'person'
ntondo	**ɳʈ**oɳ**d**o	'star'
ntambo	(mwe)**ɳʈ**ambo	'traveler'
ntembo	**ɳʈ**embo	'elephant'
ba**nt**i	ɓa**ɳʈ**i	'door'
ntangulu	**ɳʈ**aŋgulu	'basket'
vu**nd**i	vu**ɳɖ**e	'cloud'
nko**nd**e	hoɳ**d**e / qoɳ**d**e	'cultivated field'
tu**nd**a	tu**ɳɖ**a	'fruit'
ki**nd**edi	ci**ɳɖ**edi	'correct'
ku**d**anta	ku**ɗ**aɳʈo[9]	'to lie, to deceive'
nka**nd**e	haɳ**d**e / qaɳ**d**e	'food'

---

[6] Hyman (2003) provides a good overview of NC phenomena in other Bantu languages. Odden (n.d.) discusses the synchronic behavior of NC in SBK in more detail. He also mentions a contrast between syllabic nasal + C and NC as illustrated by [m̩bu], 'mosquito' vs. [mbuni], 'goat'. It is not clear whether or not this contrast has any relevance to the main point of this paper.

[7] The [o] is not a typo. Nevertheless its occurrence is unusual given that the infinitive form of most verbs in SBK and other Bantu languages ends with an [a]. Why this word ends with an [o] instead of an [a] is a puzzle. One hypothesis is that the [o] comes from Af-Maay influence since verbs in Af-Maay typically end in [o].

[8] In this form, the [m] represents a noun class marker and would hence be pronounced as a syllabic bilabial nasal, with [wa-] being the appropriate plural marker for Noun Class 1/2. This is interesting given that the consultant data shows this word belonging to Noun Class 5/6, as evidenced in the use of [ma-] for the plural.

[9] See footnote 6.

Table 6: /nt/ and /nd/ in Loan Words

SBK	Source	Gloss
hara**nt**i	Origin Unknown	'courtyard'
asa**nt**e	Standard Swahili: [asa**nt**e]	'thank you'
ɓa**nd**era	Portuguese: [ba**nd**eira] or Italian: [ba**nd**iera], possibly via Standard Swahili: [ɓa**nd**era]	'flag'
ndoni	Somali: [**ɖ**oni]	'boat'
ndeɟe	Standard Swahili: [**nd**eɟe]	'airplane'[10]
kua**nd**ika	Standard Swahili: [kua**nd**ika]	'to write'

3. A Cross-Linguistic Survey of Retroflexion

In reviewing the literature on the emergence of retroflexion, it turns out that few purely phonetic causes have been identified. The most comprehensive typological survey is perhaps Bhat (1973), which argued that most cases of retroflexion are due to external rather than internal causes. This argument is based on an examination of 150 different languages showing that this feature is geographically restricted to a handful of regions. The four regions identified include India, Australia-Southeast Asia, central Africa, and the Pacific coast of America. In addition to these four regions, three minor pockets are also identified including southern Africa, Scandinavia, and the Caucasus. The region most relevant to Kizigua is "Central Africa," which stretches from the Atlantic Ocean in the west to the Indian Ocean coast to the East. It would include the location where SBK is spoken in Southern Somalia but exclude the area where TZ is spoken along the Northeastern coast of Tanzania. The retroflexion found in these languages is described as "weak" and less prominent than in India and in Australia-Southeast Asia suggesting that it is a recent innovation in the area. These languages generally have only one retroflex stop. This makes the emergence of three in SBK in less than 200 years even more remarkable.

In all of these geographical clusters, Bhat notes that there are at least a few genetically unrelated languages. Based on this fact, Bhat postulates that retroflexion occurs in a language either "1) through inheritance from the parent language, or 2) through contact with a neighboring language that possesses the feature through 1) or 2)" (ibid.: 42). Thus, if a language did not inherit retroflexion from its parent language, the only way in which it could have developed is through contact. This, of course, raises the question of how retroflexion would have developed in the first place in the absence of neighboring languages that possess the feature. The phonetic environments identified that can induce retroflexion include (i) a preceding apical tap or trill, (ii) a following retroflexed consonant, (iii) a following back vowel, and (iv) implosion. Of these four environments, the first two can occur only in languages that already have at least one retroflex or rhotic sound. Thus, the only way a language can innovate retroflexion would be due to a following back vowel (iii) or due to implosion (iv).

With such limited phonetic environments that can induce retroflexion, it should be no surprise that they are typologically rare and that they arise more often due to contact than as an independent innovation. Bhat also postulates that these limited environments mean that there is generally a unidirectional path in which retroflex sounds develop in a language in the absence of contact. This is supported by the fact that many languages that have only one retroflex sound have a retroflex fricative, trill, or lateral. It is only at a later stage through the influence of retroflex or 'r'-like sounds that already exist in a consonant inventory that retroflex stops and affricates develop. This basically accounts for the diachronic development of retroflexion in the languages of India and Australia-Southeast Asia, which include languages with some of the largest inventory of retroflex sounds.

The two cases identified of pure phonetic motivation, a following back vowel and implosion, would actually apply to the retroflex implosive in SBK. Retroflex plosives and implosives emerging from dental or alveolar implosives is a widely attested change found in many languages of the Central

---

[10] According to my consultants, the word for 'airplane', [ndeɟe], contrasts with the word for 'bird', [ɗeɟe], which has been identified as inherited from TZ. Interestingly, the Standard Swahili word for 'airplane' simply involves semantic extension of the word for 'bird' [ndeɟe]. Since airplanes did not exist in the 19th Century, the SBK word is most likely a 20th Century loan from Standard Swahili.

African zone described by Bhat. The case of implosion basically accounts for an exception to the uni-directional path that Bhat proposed. Implosives exhibit diachronic behavior distinct from that of plosives. The result is many languages in the Central African zone with only one retroflex sound and this sound being a stop. This developmental path can be described as ɗ > ɖ̥> ɖ.

Subsequent work by Ohala (1983) and Hamann and Fuchs (2010) developed a phonetic explanation for this. According to Ohala, a key distinction between retroflex and non-retroflex sounds is that retroflex sounds involve an enlarged oral cavity behind the point of constriction. This enlarged oral cavity makes it easier to create the high transglottal pressure differential needed to maintain voicing than for alveolar stops. The change ɗ > ɖ̥> ɖ can, thus, be seen as an aerodynamic strategy to maintain voicing. Hamann and Fuchs (2010) have extended this analysis to apply to plain alveolar plosive stops as well. Citing Electromagnetic Articulography (EMMA) and Electropalatography (EPG) data from German, they show that a language currently lacking a retroflex phoneme has a voiced alveolar plosive that has a more retracted articulation than its voiceless counterpart. They argue that a voiced retroflex plosive developed directly from a voiced alveolar plosive in three unrelated languages (Dhao, Thulung, and Afar) without going through implosion as an intermediate stage. The diachronic developmental path they propose is summarized as d > ɖ.

More directly relevant to the point of this paper is Hamann & Fuch's discussion of how sociolinguistic factors are not incompatible with phonetic explanations. They show that the German speakers studied exhibit inter and intra-speaker variation in their tongue position in the production of the voiced alveolar stop phoneme. There was a greater tendency for the tongue to be retracted for [d] preceding the vowel [u] than for [i]. Different speakers also varied in the degree of retraction. While they make it clear that they are not arguing that German is on its way to developing a retroflex phoneme, they argue that the seeds for such a change are universally present in the way that [d] is articulated. There still needs to be a trigger in pushing the change to take place on a community level. Thus, they state that "sociolinguistic factors are [still] the driving force in sound changes ... A continuous updating of our pronunciation according to the *input* [my emphasis] we receive is actually happening all the time and leads to small but noticeable changes in our sound system."

Even as research on the internal motivation behind sound change continues, this body of research can still be compatible with research on external factors. Although Hamann and Fuchs (2010) do not explicitly talk about contact, this paper will show that "input" in the SBK case can be described as contact with speakers of other languages. The next section develops this argument by identifying the different groups that have historically interacted with SBK speakers. Once these groups have been identified, the relevant features of the sound systems of the languages spoken by these groups will be examined. Although there is a strong case for internal motivation in the development of /ɖ/, speakers of other languages with this sound could have still played a role by accelerating this change or at least by maintaining it. Hence, an argument will be made for external factors as primary and internal motivation as secondary.

4. The Socio-historical Contact Situation

The arrival of the Zigua in Somalia brought the Zigua in contact with a greater number of languages that contain retroflex stops. This section discusses which languages these included and the nature of contact between these groups by reviewing the history of the region.

Eno and Eno (2007) discuss how famine and drought around the turn of the 19th century served as the impetus for a migration from Tanzania to Somalia. During this time, Arab-Omani traders lured the Zigua to Somalia by promising them the opportunity to work on fertile land. Instead of being offered wage labor, however, they were tricked into slavery and forced to work on commercial plantations in Somalia. Many of them revolted and then escaped and headed southward by foot. According to estimates cited by Eno and Eno (2007), over 20,000 slaves escaped from their masters between 1865 and 1895. The fugitive slaves formed communities in a forested area in the Juba River Valley of Southern Somalia. This area became known as Gosha and a safe haven for other fugitive slaves who would later join these communities leading to the growth of these settlements. After the disestablishment of Gosha in the late 19th Century, the region became colonized by Italy until the end of World War II bringing the end of slavery. Somalia then became part of a British protectorate until it gained independence in

1960. Even with the disestablishment of Gosha, the Zigua remained in many of the same villages for almost a century. It was not until the Somali Civil War broke out in the 1990's that the Zigua were forced to move out of the region and in to UN refugee camps across the border in Kenya. From Kenya, some of the Zigua were able to resettle in their ancestral homeland in Tanzania while others resettled elsewhere including the US.

Although there has been quite a bit of political instability in southern Somalia during the past two centuries, it should be emphasized that the Zigua were generally stable and isolated from the rest of Somalia during most of this time period. The two major periods of instability would have been the time of arrival in southern Somalia and the outbreak of the Somali Civil War. These two periods would have corresponded to differences in the groups with which the Zigua were in contact. The contact situation during each of these two periods should be examined in further detail.

First of all, the port of entry for the Zigua slaves was the coastal city of Barawa (or Brava), where a Northern Swahili dialect known as Chimwiini was spoken. This is one language that should be closely examined. Yet, there were also other languages with which the Zigua came in contact in Somalia. At the time of arrival in Gosha, the Zigua encountered two major groups: (1) fugitive slaves from other ethnic groups and (2) other non-Somali groups already present in Southern Somalia at the time of arrival. The other slaves who were brought to Somalia in the 19th Century came from various ethnic groups from present-day Tanzania and Mozambique. These groups included the Shambara, the Yao, the Makua, the Ngindo, the Lima, and the Nyasa. Each of these groups spoke a different Bantu language. After a generation, however, most of them lost their native languages. Menkhaus (2003) says that the Zigua were an exception because "unlike most of the East Africans sold into captivity in Somalia, the [Zigua] were not children, but adults." Since adults are less likely to acquire other languages than children due to the critical period, the Zigua as a whole would have been less likely to shift to another language than the other groups would have been. This would have meant that children from other groups would have acquired Kizigua leading to an increase in the number of Kizigua speakers over time. The current situation is that Kizigua is no longer a language that is spoken exclusively among the Zigua people but also a language spoken by others with ancestry from different Bantu groups.

The non-Somali groups that the Zigua encountered included speakers of other Bantu languages as well as speakers of Boni (Cushitic) and Oromo (Cushitic). According to Crevatin (1993), the indigenous Bantu groups included the Bajuni, who lived on islands off the coast, and the Pokomo, who were fishers living in small villages near the coastal city of Kisimayo. Many details about the nature of the interaction between the Zigua and these various groups are unclear. According to oral reports, however, the Zigua absorbed themselves into existing Bantu communities when they arrived in Southern Somalia (Eno and Eno 2007). Their relationship with other Bantu groups appears to have been more positive, however, than their relationship with ethnic Somali from the North. Even after the establishment of Gosha, the Zigua frequently had to defend themselves from attack from Somali nomads. Aided in part by firearms provided by the Sultanate of Zanzibar, they were generally quite successful in defending themselves (Menkhaus 2003). The Somali Civil War in the 1990's, of course, changed the situation forcing a second diasporic migration (Eno and Eno 2007).

A common theme that emerges in all accounts of interaction with the Bantu groups in Southern Somalia is that of absorption. Interviews with SBK speakers confirm that intermarriage took place between different Bantu groups, though intermarriage with ethnic Somalis has historically been less common and frowned upon. For example, one speaker traces his ancestry to both the Makua tribe from Mozambique and to the Zigua tribe. The frequency of intermarriage between different Bantu groups, however, is uncertain. Nevertheless, this same speaker also said that the Zigua were historically more numerous and that this would explain why many of these other languages such as Makua are no longer spoken in Southern Somalia. So if absorption into the indigenous Bantu groups also took place, their numbers may have also been historically too few for languages such as Chimwiini, Bajuni, or Pokomo to have taken over. There still, however, may be substrate features from these languages in the SBK spoken today.

The second major period of instability began at the outbreak of the Somali Civil War. This appears to have been a turning point with major sociolinguistic consequences. All of the speakers interviewed grew up during this time period and are proficient in varying degrees in Somali, Af-Maay,

Standard Kenyan Swahili, and English. This contrasts with the pre-Civil War situation in which profi-
ciency in Somali was far less common. During the 1980's, for example, Crevatin (1993) mentions that
in many Zigua villages, it was quite common to find people who had minimal or no knowledge of So-
mali. Crevatin also provides further evidence of isolation from the Somali language by mentioning the
lack of loanwords. The four speakers interviewed for the current project, however, revealed a very
different situation for a younger generation. All of these speakers expressed conscious awareness of
generational difference in how their language is spoken. They also reported increased usage of both
Somali and English loanwords compared to older speakers. The implication of this generational shift is
that any contact-induced changes that could be attributed to Somali would more likely be recent
changes than changes that took place soon after arrival in Somalia. We have evidence for retroflexion,
however, in Crevatin's (1993) data from the 1980's. Thus, it appears that retroflexion predates the Civ-
il War and may have been more likely due to the influence of languages other than Somali.

Aside from Kizigua, Af-Maay is the other language that is widely spoken within the Somali Bantu
community. It is genetically classified as a Cushitic language, but its historical status within the Somali
Bantu community remains uncertain. For example, while all speakers interviewed for the current pro-
ject speak Af-Maay as a second language, no material has been found indicating when the Zigua first
started to speak Af-Maay. Thus, it is not known if contact with Af-Maay is a recent phenomenon as is
the case with Standard (Northern) Somali or a historic one. Its relationship to Somali is also uncertain.
According to Paster (2006), whether it is a dialect of Somali or an independent language is an unre-
solved question.

The next section will take a closer look at the sound systems of the languages mentioned in this
section.

5. Possible External Sources of Retroflexion

As discussed in Section 4, the languages that came in contact with Kizigua in Somalia include
both Cushitic and Bantu languages. Relevant information about the coronal stop inventories for
languages with available data is presented in Table 7 below.

Table 7: Retroflex Inventory of Contact Languages

Language	Family	ɳ t̺	ɳ ɖ	Coronal Implosives[11]	Source
Maay	Cushitic	--	--	ɗ	Paster (2006)
Boni (Aweera)	Cushitic	--	--	ɗ	Nurse (1985)
Oromo	Cushitic	--	--	ɗ̺ (depends on dialect)	Gragg et al. (1982)
Standard Somali	Cushitic	--	--	ɖ̺ (possibly ɗ̺ in the 19th century)	Saeed (1999)
Northern Swahili – Bajuni Dialect	Bantu	--	✓	--	Nurse (1985), Nurse and Hinnebusch (1993)
Northern Swahili – Chimwiini Dialect	Bantu	✓	✓	(ɗ)	Nurse (1985), Nurse and Hinnebusch (1993)
Pokomo	Bantu	✓	✓	(ɗ)	Nurse (1985), Nurse and Hinnebusch (1993)

Among the languages listed above, only two have been unambiguously described as having a
retroflex implosive. One of these is Oromo, a language described in the previous section as one with
which there is a known history of contact with Kizigua. Even so, not all dialects have been described
as having /ɗ̺/. The other language is Somali, but again this appears to vary by dialect. Saeed (1999)
mentions that while the modern Standard variety has a retroflex plosive, it may have historically been

---

[11] The parentheses in this column indicate a loan phoneme.

an implosive in the 19th Century. So if /ɗ/ in SBK were borrowed from Somali, it would have been borrowed in the 19th Century. This, however, would contradict the socio-historical evidence suggesting minimal interaction with Somali until recent times. Oromo, however, also has /ɗ/ as well as socio-historical evidence of contact during the 19th Century. Thus, even though there is linguistic evidence showing Somali to be a possible source of /ɗ/, there is stronger evidence that Oromo could have been the source based on *both* the comparative linguistic data and the socio-historical data.

Yet, there could be other possibilities. Going back to Section 3, it was mentioned that ɗ > ɖ is an attested internally motivated sound change in other languages. So if we look beyond descriptions that explicitly mention a retroflex implosive, we can see that just about all of the other languages in the list above have /ɗ/, though in some cases such as in Pokomo and Chimwiini it is a loan phoneme. It seems quite possible that SBK may have borrowed /ɗ/ from any of these languages and that [ɗ] subsequently developed as an allophonic variant. It also seems possible that /ɗ/ may have actually been post-alveolar or retroflex in these languages all along even if not explicitly described as such due to the lack of a contrast.

Another major pattern shown above is a language family split that also corresponds to whether or not a language has retroflex NC. This is a feature that is limited to Bantu languages and completely absent in Cushitic languages. While NC sequneces do not appear to be phonotactically prohibited in Cushitic, it does not seem that such NC sequences ever function as single phonemic units as is the case for many Bantu languages. Likewise, NC never occurs in word-initial position in Cushitic while many words in SBK do begin with NC. Interestingly, the SBK word [ndoni], 'boat' is a word likely borrowed from Somali or another Cushitic language. In Somali, however, the word is [ḍoni] and begins with a dental stop instead of a prenasalized one as in SBK. What seems clear is that if retroflex NC in SBK were borrowed, it would have been much more likely borrowed from a Bantu language than from a Cushitic language because Bantu languages have a very similar phonological unit not found in Cushitic languages.

One caveat to note is that not all of the sources for these languages describe varieties spoken in Somalia in the 19th Century. The data from Northern Swahili dialects and Pokomo, for example, comes from varieties spoken in Northern Kenya. If SBK borrowed from these languages, SBK would have borrowed from Somali varieties of these languages spoken during the 19th Century. Yet, if we assume that these documented varieties are similar enough, we can clearly see that there are multiple languages that have the three retroflex stops in question and that any of these languages could have been the source of these sounds as well as undocumented varieties of them. The next section will develop an analysis of how they might have been transferred to SBK.

6. Analysis of the Development of Each Retroflex Sound

The discussion of the diachronic development of all three retroflex stops will be divided into two subsections below. The first subsection is on the retroflex implosive while the second subsection is on the prenasalized stops. This division reflects two different stories that will be presented. The first one about the implosives offers more evidence for an internally motivated explanation while the second one offers less evidence for internal motivation. In both of these cases, however, external motivation will be described as the primary driving force for change. Thus, while there is evidence for internal motivation for /ɗ/, an argument will be made that this does not preclude the role of contact as the initial trigger. Hence, contact is argued to provide a unifying explanation for the development of all three stops.

6.1. The Retroflex Implosive, ɗ

As discussed in Section 2, [ɗ] and [ɗ] occur in complementary distribution with the retroflex occuring before back vowels, /u, o, a/, and the alveolar occuring before front vowels, /i, e/. Two of the environments described by Bhat (1973) as inducing retroflexion cross-linguistically are applicable here namely implosion and a following back vowel. Since retroflexion involves retraction of the tongue body, its co-occurence with back vowels would be naturally compatible with the tongue position involved (Bhat 1974). Even in a language lacking retroflex phonemes like German, Hamann and Fuchs

(2010) showed that the tongue position is more retracted and retroflex preceding /u/ than when preceding /i/. Thus, there are clear phonetic reasons motivating this pattern due to co-articulation with the following vowel.

Further evidence for internal motivation can be found in the presence of at least one loan word containing [ɖ]. The one identifiable loanword in the collected data that contains a retroflex implosive is [ɖunia], which comes from Swahili (and most likely a Kenyan variety), a language that has not been described as having this sound. Evidence that this is a loan comes from its absence in dictionaries of TZ (Kisbey 1906; Mochiwa 2008) and the fact that speakers consciously describe this word as a loan from Swahili that is not used by older speakers. If [ɖunia] were a loan from Swahili, the question to ask would be why not [ɗunia] as in Swahili? The fact that a retroflex articulation occurs instead of an alveolar articulation suggests that this word has become nativized into the SBK phonological system.

Having established evidence for internal motivation for the development of /ɖ/, the next question to ask would be if evidence can be established against external factors. In order to completely rule out external factors, evidence that /ɖ/ developed in TZ prior to migration to Somalia as well as evidence showing the absence of neighboring languages that contain /ɖ/ would need to be available. No known language from the Northeastern coast of Tanzania, however, has been identified as having a [ɖ] either in the 19th Century or in the present-day. Furthermore, we know that Oromo has /ɖ/ and that Bantu languages spoken in Southern Somalia and Boni contain /ɗ/, which may have subsequently developed into /ɖ/. With a greater number of languages in this region with /ɖ/ or /ɗ/, it seems much more likely that /ɖ/ in SBK developed through contact with one or more of these languages. Thus, more exposure to retroflex implosives would have been possible in Somalia than in Tanzania.

Even if a change from alveolar to retroflex began in Tanzania, the change could have been accelerated once Kizigua came in contact with one of these languages. The change could have started in certain contexts such as preceding the vowel /u/ and then it could have gradually spread to preceding the vowel /o/ and then preceding the vowel /a/ over the course of multiple generations. This would be compatible with Thomason and Kaufman's (1988) claim about the primacy of social factors in language change. Still, more information about the nature and extent of contact with these groups would be helpful in order to develop a better understanding of the mechanisms of how contact could have played a role in the emergence of the retroflex implosive in SBK. For the time being, the fact that there was known contact with speakers of languages containing this sound should not be seen as trivial. Even with evidence in support of internal factors, there is insufficient evidence to completely rule out external factors.

6.2. Retroflex NC Stops: /ɳɖ/ and /ɳʈ/

For both of the retroflex NC complexes in SBK, evidence in support for internal phonetic motivation in their development is more lacking than for the retroflex implosive. First of all, there are too few cases to know of any phonetic reason as to why alveolar NC or prenasalized stops in general would diachronically retract. None of the phonetic environments described by Bhat (1973) would be applicable for NC. The NC stops do not involve implosion nor is their occurrence restricted to certain vocalic environments. Both front and back vowels can follow retroflex NC as illustrated by the words [ɓaɳʈi] and [ɳʈoɳɖo]. Furthermore, internal motivation would not explain why an alveolar/retroflex contrast would even develop in the first place. Hamann and Fuchs (2010) does not seem applicable either since their explanation of how a voiced retroflex stop can develop is based on a difference between voiced and voiceless stops. In the SBK case, both voiced and voiceless retroflex NC stops are present.

One of the few cases of retroflexion of alveolar NC is described by Nurse and Hinnebusch (1993), which mentions *nt and *nd retroflexion in various Northeastern Bantu languages as motivated by perceptual reasons due to the difficulty of distinguishing between dental and alveolar sounds. The dental sounds in these languages have been described as having entered the language from borrowing of vocabulary from Cushitic languages (Nurse 1985). Yet, even in this case, this would have been a contact-induced change that triggered a subsequent change rather than a purely internally motivated one.

Given the fact that there is evidence of social interaction with speakers of languages in Southern Somalia that have retroflex NC including various dialects of Northern Swahili, the most likely explanation is that these sounds were borrowed from one or more of these languages rather than innovated

independently. There are several possibilities in terms of how /ɳɖ/ and /ɳʈ/ might have entered Kizigua through contact. To illustrate how, Table 8 below includes a list of words with retroflex NC in SBK along with corresponding forms in TZ and in Chimwiini. Although, it is possible that other Bantu languages may have also been involved, Chimwiini lexical items appear to have a much better match than other languages with SBK lexical items.

Table 8: Corresponding Forms in Chimwiini

TZ	SBK	Chimwiini[12]	Gloss
-nkundu	-huɳɖu	-hu:ɳɖu	'red'
nkonde	hoɳɖe	hoɳɖe	'cultivated field'
kenda	ceɳɖa	keɳɖa	'9'
tunda	tuɳɖa	tuɳɖa	'fruit' (sg.)
kintu	ciɳʈu	ciɳʈu	'thing'
m̩ṇʈu	m̩ṇʈu	muɳʈu	'person'
ntembo	ɳʈembo	(te:mbo)	'elephant'
ntondo	ɳʈoɳɖo	(no:ta)	'star'

One possibility is that Kizigua speakers borrowed all words containing /ɳɖ/ and /ɳʈ/ in Chimwiini or another Bantu Language spoken in Somalia. In fact, we can see a very good match between some of the forms in Chimwiini with their corresponding forms in SBK. For example, [tuɳɖa] in SBK is exactly the same as in Chimwiini but differs from the TZ form only in having a retroflex NC rather than an alveolar one. In some cases such as TZ [kintu] → SBK [ciɳʈu], there are two sound changes that result in the SBK form looking identical to the Chimwiini form. Because of this similarity, the outcome of borrowing of words with retroflex NC from Chimwiini or another Somali Bantu language would resemble a sound change involving a change from alveolar NC to retroflex NC. It could even account for changes that are much more typologically common such as palatalization of /k/ since the borrowing of [ciɳʈu] would also give the appearance of k → c. Thus, it appears that borrowing in the case of related languages can result in changes that resemble sound changes due to internal phonetic motivation.

Table 8 also presents data suggesting that even if borrowing was involved for a few cases, it may not have been the complete story. Not all words identified in SBK that contain retroflexes have been identified as having corresponding cognates in Chimwiini or another Bantu language. For example, the word for 'star' in TZ lacks a corresponding form in Chimwiini. Also, the word for 'elephant' in Chimwiini begins with a plain /t/ while the SBK form begins with a retroflex NC. One possibility to account for these particular words would be that borrowing of a handful of words from one of these languages triggered subsequent changes in the form of other lexical items. In other words, what may have initially began as a change in lexical form in a few words may have spread to other words through substitution of all alveolar NC with retroflex NC. The outcome of this contact situation would be a set of sound correspondences between SBK and TZ that appear to be a result of internal motivation.

Yet, there is another possibility. Given that intermarriage occurred between members of different Bantu groups, it would also be worth considering the possibility that it was native speakers of other Bantu languages rather than native Kizigua speakers who were the agents of change. When acquiring Kizigua, these other Bantu speakers may have simply pronounced all words with alveolar NC in TZ with retroflex NC in Kizigua as found in these languages. They would have then passed this pronunciation along to future generations of Kizigua speakers. This explanation would be able to account for why some words with retroflex NC in Kizigua do not have corresponding forms in Chimwiini or other Somali Bantu languages. If it turns out to be a case of second language acquisition for non-native speakers of Kizigua, there would have been an across-the-board substitution of sounds in Kizigua with similar forms in Chimwiini or another Somali Bantu language. This also seems to be a plausible sce-

---

[12] Data comes from Nurse and Hinnebusch (1993) and Kisseberth and Abasheikh (2004). The original transcriptions have been converted to better illustrate the match between Chimwiini and SBK. The original transcriptions used a superscript [ʳ] to represent retroflexion as in [ndʳ].

nario given the fact that multilingualism and intermarriage were both part of the social history of the region. Furthermore, this would also explain why the across-the-board change of alveolar NC to retroflex NC looks like an internally motivated sound change.

To summarize, there is more than one possible contact-induced explanation for how retroflex NC developed. Borrowing of a few lexical items from another Somali Bantu language is one possibility. This may have triggered changes in the rest of the lexicon leading to the sound correspondence we see today between alveolar NC in TZ and retroflex NC in SBK for inherited vocabulary. Another explanation is based on native speakers of other Somali Bantu languages learning Kizigua and substituting instances of alveolar NC in TZ with the retroflex form found in their native language. This pronunciation would have then spread across the community and would have been passed down to subsequent generations. Evidence for internal motivation, on the other hand, seems lacking. There does not seem to be a clear explanation as to how internal motivation would lead to the development of both a voiceless and a voiced retroflex NC stop.

7. Conclusion

This paper began by introducing a "rediscovered language" (Crevatin 1993) isolated from other speakers of the language in Tanzania for more than a century. This isolation was created in the 19th Century when people from various Bantu tribes from the coastal areas of modern-day Tanzania and Mozambique were brought to Somalia to work as slaves. Since this historic migration, the Kizigua language as spoken in Somalia has developed a consonant inventory with three retroflex phonemes. Previous cross-linguistic research on the emergence of retroflexion was reviewed. It showed that there are a limited number of phonetic environments that can induce retroflexion. Most cases of the emergence of retroflexion are due to contact. The socio-historical context of the Kizigua language was reviewed. It suggested several possible languages to examine as potential sources of the retroflex stops. Ultimately, it was concluded that an explanation based on contact could account for the emergence of all three retroflex stops. This explanation, however, does not preclude the secondary role of internal motivation for the emergence of ɖ. Thus, it does not seem to be simply a coincidence that the Northeastern Coast of Tanzania is a region that lacks languages with retroflex sounds while Southern Somalia has more languages with retroflexion. Overall, the data and analysis presented in this paper support Thomason and Kaufman's (1988) argument for the primary role played by external factors in language change.

Future research could further investigate the phonetic properties of all three retroflex sounds. Can we learn something about the phonetic properties of these under researched sounds that may better explain their development? Secondly, more research needs to be done on the history of contact with different groups in Southern Somalia. More data on these languages would also be helpful. There are still many gaps in the historic record and in the linguistic data that could shed much light about what happened. For the time being, this paper has presented the basic facts about an unusual case of the diachronic emergence of retroflexion as well as an under-researched case of language change in the context of a history of migration.

References

BARLEW, JEFFERSON. 2013. Point of View in Mushunguli Locatives. *Selected Proceedings of the 43rd Annual Conference on African Linguistics*, ed. by Olanike Ola Orie and Karen W. Sanders, 115–129. Somerville, MA: Cascadilla Proceedings Project.

BESTEMAN, CATHERINE. 2012. Translating Race across Time and Space: The Creation of Somali Bantu Ethnicity. *Identities: Global Studies in Culture and Power* 19.285–302. doi:10.1080/1070289X.2012.681862.

BHAT, D. N. S. 1973. Retroflexion: An Areal Feature. *Working Papers on Language Universals* 13.22–67.

BHAT, D. N. S. 1974. Retroflexion and retraction. *Journal of Phonetics* 2.233–237.

BRENZINGER, MATTHIAS. 1987. East African Beekeeping Vocabularies: kiZigua. *Afrikanistische Arbeitspapiere* 9.113–123.

CREVATIN, FRANCO. 1993. Incontri con l'Africa. *Studi italiani di linguistica teorica e applicata* 22.11–26.

ENO, OMAR; and MOHAMED ENO. 2007. The Journey Back to the Ancestral Homeland: the Return of the Somali Bantu (Wazigwa) to Modern Tanzania. *From Mogadishu to Dixon: the Somali Diaspora in a Global Context*, ed. by Abdi Kusow and Stephanie R. Bjork, 13–43. Trenton, NJ: Red Sea Press.

GRAGG, GENE B.; and WITH THE HELP OF TERFA KUMSA AND OTHER OROMOS. 1982. *Oromo dictionary*. East Lansing, Mich.: African Studies Center, Michigan State University, in cooperation with Oriental Institute, University of Chicago.

GUTHRIE, MALCOLM. 1967. *Comparative Bantu : an introduction to the comparative linguistics and prehistory of the Bantu languages.* 4 vols. Farnborough: Gregg Press.

HAMANN, SILKE; and SUSANNE FUCHS. 2010. Retroflexion of Voiced Stops: Data from Dhao, Thulung, Afar and German. *Language and Speech* 53.181–216. doi:10.1177/0023830909357159.

HYMAN, LARRY. 2003. Segmental Phonology. *The Bantu Languages*, ed. by Derek Nurse and Gérard Philippson, 42–58. Routledge.

KENSTOWICZ, MICHAEL. 1988. Tone and Accent in Kizigua-a Bantu Language. *Certamen Phonologicum*, ed. by Pier Marco Bertinetto and Michele Loporcaro, 177–188. Turin: Rosenberg & Sellier.

KISBEY, WALTER HENRY. 1906. *Zigula-English dictionary. Compiled for the Universities' Mission to Central Africa by Rev. Walter H. Kisbey*. England.

KISSEBERTH, CHARLES W; and MOHAMMAD IMAM ABASHEIKH. 2004. *The Chimwiini lexicon exemplified*. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

LEWIS, M. PAUL, GARY F. SIMONS, and CHARLES D. FENNIG (eds.) 2013. *Ethnologue: Languages of the World*. 17th ed. Dallas, Texas: SIL International. http://www.ethnologue.com/.

MAHO, JOUNI FILIP. 2009. *NUGL Online: The online version of the New Updated Guthrie List, a referential classification of the Bantu languages*. http://goto.glocalnet.net/mahopapers/nuglonline.pdf.

MARTINEZ, MICHAL. 2013. Documenting Refugee Languages in the Diaspora: The Boise Language Project. http://works.bepress.com/michal_martinez/17.

MENKHAUS, KEN. 2003. Bantu ethnic identities in Somalia. *Annales d'Ethiopie* 19.323–339.

MOCHIWA, Z. S. M. 2008. *Kizigula : msamiati wa Kizigula-Kiswahili-Kiingereza = Zigula-Swahili-English lexicon*. Dar-es-Salaam: Languages of Tanzania Project University of Dar-es-Salaam.

NURSE, DEREK. 1985. Dentality, Areal Features, and Phonological Change in Northeastern Bantu. *Studies in African Linguistics* 16.243–279.

NURSE, DEREK. 2010. The Decline of Bantu in Somalia. *Essais de typologie et de linguistique générale. Mélanges offerts à Denis Creissels*, ed. by F. Floricic, 187–200. Lyon: ENS Editions.

NURSE, DEREK; and THOMAS J. HINNEBUSCH. 1993. *Swahili and Sabaki: a linguistic history*. University of California publications in linguistics v. 121. Berkeley: University of Los Angeles Press.

ODDEN, DAVID. n.d. . *Mushunguli - Chizigua language of Somalia*. http://www.ling.ohio-state.edu/~odden/mushunguli/.

OHALA, JOHN. 1983. The Origin of Sound Patterns in Vocal Tract Constraints. *The Production of Speech*, ed. by P. F. Macneilage, 189–216. New York: Springer.

PASTER, MARY. 2006. Aspects of Maay Phonology and Morphology. *Studies in African Linguistics* 35.73–120.

SAEED, JOHN I. 1999. *Somali*. Amsterdam: J. Benjamins Pub. Co.

SAMARIN, WILLIAM J. 1967. *Field linguistics: a guide to linguistic field work*. New York: Holt, Rinehart and Winston.

THOMASON, SARAH GREY; and TERRENCE KAUFMAN. 1988. *Language contact, creolization, and genetic linguistics*. University of California Press.

TSE, HOLMAN. 2013. Methodological considerations in the study of sociophonetic variation in an underdocumented minority language: Somali Bantu Kizigua as a case study. *Proceedings of the 4th Edition of Journées d'Études Toulousaines (JéTou)*, 129–139. University of Toulouse. http://jetou2013.free.fr/documents/JeTou2013-Actes-p129-139-Tse.pdf.

Selected Proceedings of the 44th Annual Conference on African Linguistics

edited by Ruth Kramer, Elizabeth C. Zsiga, and One Tlale Boyer

Cascadilla Proceedings Project   Somerville, MA   2015

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Tse, Holman. 2015. The Diachronic Emergence of Retroflexion in Somali Bantu Kizigua: Internal Motivation or Contact-Induced Change? In *Selected Proceedings of the 44th Annual Conference on African Linguistics*, ed. Ruth Kramer et al., 277-289. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3145.