Kandidatuppsats i Lingvistik
Intstitutionen för Filosofi, Lingvistik och Vetenskapsteori
Göteborgs Universitet
Handledare: Eva-Marie Ström
Examinator: Malin Petzell
Student: Victor Bogren Svensson

GÖTEBORGS UNIVERSITET

FILOSOFI, LINGVISTIK OCH VETENSKAPSTEORI

A preliminary analysis of the phoneme inventory of Ulubuya, a variety within the Masaba dialect continuum

Table of Contents

1. Introduction

Ulubuya is a variety within the Masaba dialect continuum, which in turn is classified under the Masaba-Luhya group. Ulubuya is spoken in eastern Uganda, close to the Kenyan border. At the time of this study, Ulubuya was an undocumented variety. I took interest in the variety when I came into contact with a native speaker during a course in linguistic fieldwork at the University of Gothenburg. I then decided to make a contribution to the study of the Masaba dialect continuum and hopefully this study will contribute to our knowledge of Ulubuya as well as help protect and preserve the variety.

This study shows that Ulubuya has five vowels, both long and short. The consonant inventory is relatively limited and shows some resemblance to the reconstructed inventory for Proto-Bantu. It exhibits voiceless plosives without voiced counterparts and has a comparatively large set of fricatives and affricates, the vast majority being unvoiced. Ulubuya seems to have undergone a historical process of spirantisation. There are important morpho-phonological processes in Ulubuya, resulting in nasal-consonant sequences, glide formation and glide hardening. The language data was obtained during interviews with a native speaker[1] of Ulubuya.

2. Aim and Purpose

The purpose of this essay is to make a preliminary sketch of the phoneme inventory of Ulubuya. The primary aims are to describe and analyse the segmental phonology of Ulubuya as well as its syllable structure. Furthermore, the scope of this study also includes a preliminary analysis and description of several phonological processes, both diachronic and synchronic, in Ulubuya. These include glide formation, spirantisation and the phonological influence nasals exert on the realisation of other consonants. This analysis and description of Ulubuya is supported by phonetic data. The study is preliminary but contributes to our knowledge of a so far undocumented language variety.

1 In accordance with the terminology employed by Chelliah (2010), the term “native speaker” will be used to denote the speaker of Ulubuya who was interviewed while gathering data about Ulubuya, to avoid using any potentially controversial terms such as informant.

3. Method

3.1. Theoretical Framework

This study seeks to describe and analyse the phoneme inventory of Ulubuya from a typological perspective. The data is analysed in comparison to common phonological features among Bantu languages as described primarily by Hyman (2003), Maddieson (2003) and Kisseberth & Odden (2003). Ulubuya is also described in light of the proposed reconstructed phonemes of Proto-Bantu as described by Meeussen (1967) and the proposed lexical reconstructions of Proto-Bantu. Moreover, the study takes a wider typological perspective beyond the Bantu language family and analyses and describes the data in relation to general phonological typology.

3.2. Native Speaker

The native speaker is called Kato Francis[2] and comes from eastern Uganda. He was born in 1987. Ulubuya is the primary language spoken in his native village. In 2006 he moved to Kampala, the capital of Uganda, where mainly Luganda is spoken. Excluding Ulubuya, which he regards as his strongest language, he also speaks Luganda and English, the latter of the two he learnt primarily in school. Kato Francis has a high level of education and was at the time of the interviews studying at postgraduate level at a Swedish university.

3.3. Interviews

This section deals with the details of the interviews during which the language data was collected. It also discusses difficulties that were encountered and strategies used to overcome them.

The interviews were conducted at the University of Gothenburg approximately from the beginning of February to the end of March, 2014. In total, six interviews were conducted at approximately one and a half hours each. Five of these interviews took place during a course in linguistic fieldwork at the university. One session was conducted separately by me and two other students at the university who were also writing papers on Ulubuya. Only a very rough

---

2 The native speaker has given his permission to be mentioned by name in this paper. Without his contribution and patient cooperation this study could never have been complete and I wish to extend my gratitude.

transcription of the data was done during the university course and no attempt was made to provide a complete description of the phoneme inventory.

All the data obtained during the sessions were recorded. While recording, the native speaker was asked to repeat all words and phrases twice. Before recording the target language, the words and phrases were repeated by the interviewer in order to confirm with the native speaker that the data was correctly transcribed.

After the sessions the collected data was further studied and analysed using the program Praat in order to make sure that the phonetic transcription made during the interviews were accurate. This was also done so that the data could be studied acoustically. This was a very important part of the study, particularly when studying vowel formants and length as well as tones.

The primary method used during the interviews was elicitation. There are several kinds of elicitation (Bowern 2008) and the one primarily used during the interviews was having the native speaker provide translations of words, phrases and sentences from a contact language (English) to the target language (Ulubuya). The choice of words was primarily based on the Swadesh list (Chelliah 2010).

Using elicitation has several drawbacks as the contact language might influence the data elicited from the target language. A translation into the target language that is too literal might be nonsensical and if the translation is not sufficiently literal one might fail to elicit the appropriate data. Elicitation through a target language might also cause the data to be unnatural as one might elicit phrases and sentences that are rarely or never used in the target language (Bowern 2008). Since elicitation was the primary method used, these difficulties were all encountered during the interviews. Strategies to avoid these difficulties were employed, for example by giving the native speaker more control of the sessions. This was achieved by, for instance, asking the native speaker to freely provide phrases used during greetings as well as naming objects in a certain semantic field, such as words for the different relationships within a family.

Occasionally the native speaker had difficulties recalling certain words that were attempted to be elicited. This was remedied on the native speaker's own initiative as he in between sessions, via email, consulted family members living in Uganda for words that he was unable to recall.

Chelliah (2010) claims that a list of 500-700 words ought to be sufficient for the purpose of developing the entire phoneme inventory of a given language. If the phoneme inventory of a closely related language is known, a smaller list might be sufficient. For this study approximately 250 words were used. Even though the amount of words used were less than what is recommended by Chelliah (2010), the phonology of Bukusu (Mutonyi 2000) as well as other dialects within the the Masaba dialect continuum (Siertsema 1981) was known. The amount of data used for this study can be considered to satisfy the criteria laid down by Chelliah (2010).

Chelliah (2010) recommends that for phonological fieldwork one should start off with data from a single speaker and then move on to compare the result with other native speakers. One might thus criticize this study for not basing the data on a sufficient amount of native speakers since only one speaker was interviewed. Further criticism might be directed towards the fact that the native speaker is not currently living in an environment where the target language is spoken, which might have a negative effect on the quality of the language data obtained.

The native speaker, however, is young and free from speech impediments, thus making him fulfilling several of the criteria laid down by Chelliah (2010) regarding what constitutes a native speaker appropriate for phonological investigations. Furthermore, being isolated from ones language community for extended durations of time primarily affects a speaker's ability to recollect words and phrases in the language. It seems to have a minimal effect on a speaker's ability to recollect and produce the sounds of the language. In addition, some strategies were employed to overcome these shortcomings, such as conducting the interviews without long time spans in between as well as starting the sessions with the greetings used in Ulubuya. Both these strategies can help the native speaker to recall and use the variety being studied. Finally, due to the fact that the variety was at the time of this study undocumented, this study still provides valuable knowledge both to the field of linguistics and speakers of Ulubuya.

4. Background

4.1. Classification of Ulubuya

In this section I describe the classification of Ulubuya based on different sources. They are generally in agreement regarding the classification, although different sources employ different terminologies.

Bantu languages are a subgroup of the Niger-Congo language family (Andersson 2001). Traditionally Bantu languages have been classified in various zones under the system established by Guthrie (1967). In Guthrie's system however, languages are classified based on geographic distribution rather than based on their diachronic relationship to other languages (Maho 2003). Guthrie makes no mention of Ulubuya, however, the Masaba-Luhya group is classified as E30 (Guthrie 1967).

Maho (2003) provides an updated version of Guthrie's classification. In this updated version, the Masaba-Luhya group is to be found. The group is classified as E30. Masaba is classified as E31 and has seven variations/dialects, one of which is (Ulu-)Buya, known as E31F. Ulubuya is thus classified as belonging to the Masaba dialect continuum (Maho 2003).

According to Glottolog (2014), Ulubuya is classified as a dialect of Masaba. Masaba is spoken in the eastern parts of Uganda, near the Kenyan border. It is classified by Glottolog under what is known as "Great Lakes Bantu". Masaba is also included in Ethnologue's (2014) database, which also includes Gisu as an alternate name for Masaba. Ulubuya is classified as one of the dialects under Masaba, which in turn is classified under the Masaba-Luhya group (E31).

According to Maho (2003) Bukusu is, just like Ulubuya, a variety within the Masaba dialect continuum. Mutonyi (2000) however, refers to Bukusu as a language rather than a dialect and claims that it is spoken in western Kenya. Thus, the status of Bukusu is not certain but regardless of what status it is given it is nonetheless closely related to Ulubuya and therefore relevant to this study.

In one of the very few publications on Masaba, Siertsema (1981) refers to native speakers from the area where Ulubuya is spoken. In another publication, Siertsema (1968) also writes about the various dialects spoken in the area where Kato Francis comes from but makes no reference to or mention of Ulubuya in any of the publications. Since the names of the variants are not standardised, it is possible that she might have discussed Ulubuya as well, in addition to the other variants.

In conclusion, Ulubuya is a part of the Masaba (known as E31) dialect continuum but its relationship to other dialects is not currently known. The sources cited generally agree to the classification of the variant. Ulubuya is spoken in the eastern parts of Uganda, nearby the border to Kenya.

4.2. Proto-Bantu Phonology

In this section I write about the reconstructed phoneme inventory of Proto-Bantu as well at its syllable structure. I also discuss some controversies and different suggestions for the reconstructed phonemes.

The phoneme inventory of Ulubuya is compared to Proto-Bantu in this study. Therefore a brief presentation of the reconstructed phoneme inventory of Proto-Bantu will follow, which will be contrasted with Ulubuya in order to uncover diachronic changes in the phoneme inventories. The proposed phoneme inventory of Proto-Bantu has a relatively small set of consonants and vowels. The proposed phonemes presented in (1) and (2) are taken from Meeussen (1967).

(1) Proto-Bantu Consonants

*[3]m	*n	*ɲ	
*b	*d	*j	*g
*p	*t	*c	*k

(2) Proto-Bantu Vowels

*i̧		*u̧
*i		*u
*e		*o
	*a	

Hyman (2003) claims that there is some disagreement regarding the reconstructed consonants, in particular the series of voiced plosives. The disagreement concerns whether or not they ought to be reconstructed as a series of continuants (/β/, /l/, /ɣ/) rather than plosives. There is a similar disagreement regarding /c/ and /j/, whether they should be regarded as affricates or

3 *= reconstruction

plosives, or even if they should be regarded as palatals, considering that they are realised as /s/ and /z/, respectively, in many modern day Bantu languages (Hyman 2003).

Regarding syllable structure, Proto-Bantu had a very limited set of permitted structures consisting of CV, CVV and V, N. Depending upon how the nasal + consonant feature is analysed, it could constitute the only consonant cluster in Proto-Bantu, provided it is analysed as two separate segments and not as a single pre-nasalised consonant (Hyman 2003).

4.3. Bukusu Phonology

Here I describe the phoneme inventory of Bukusu as well as some phonological processes within the language. Bukusu is closely related to Ulubuya and is therefore relevant to this study. Bukusu is spoken in western Kenya (bordering the area where Ulubuya is spoken) and is classified under Guthrie's system as E31c (Mutonyi 2000).

Bukusu has a set of five vowels, similar to many other Bantu languages. Long and short vowels are distinguished but these are according to Mutonyi (2000) not separate phonemes and length should be treated as a supra-segmental unit. This is in contrast with the classification employed by Austen (1974) and Blois (1975), which regard them as separate phonemes. Bukusu exhibits glide formation for high vowels /i/ and /u/ into /y/[4] and /w/, respectively and also has vowel harmony for its prefixes and suffixes (Mutonyi 2000). The phonemes presented in (3) and (4) are taken from Mutonyi (2000).

4 /y/ is used by Mutonyi (2000) to represent the IPA symbol /j/ and when referring to Bukusu in this paper I have adopted the praxis used by Mutonyi (2000).

(3) Bukusu vowels

i	u	ii	uu
e	o	ee	oo
a		aa	

(4) Bukusu Consonants

p	t	č	k	Stops
(mb)[5]	(nd)	(ɲj)	(ŋg)	
f	s	x		Fricatives
β				
m	n	ɲ	ŋ	Nasals
l r				Liquids
w		y		Glides

5. Result and analysis of the data

5.1. Introduction and presentation of the data

This section deals with how the language data is presented in this paper in addition to a brief presentation on the conventions employed when presenting segmented underlying forms.

The International Phonetic Alphabet (IPA) is used in this phonological description of Ulubuya. The data is presented first with a phonological transcription, followed by a proposed segmented underlying form, then a translation into English and finally further details about the language data, such as noun class.

The language data presented in this study is also marked for surface tone, based on consultation with the native speaker, listening to the language data as well as using Praat to study the data acoustically. High tone is marked according to IPA praxis and low tone is left unmarked. It is beyond the scope of this study to provide an exhaustive analysis and description for the phonological function of tone in Ulubuya and it is included here primarily to serve as reference for potential future studies.

5 The pre-nasalised consonants are provided in brackets by Mutonyi (2000) due to their limited distribution.

In the presentation of the data, C_1 refers to the stem initial consonant and C_2 refers to the stem internal consonant. In addition, long vowels have been analysed as underlying short if they precede a nasal-consonant sequence, follow derived glides and occur within prefix haplology. See section 5.3. for discussion.

5.2. Consonants

In this section you will find a summary of the consonant inventory in Ulubuya. Each of subsections deal with distribution of the consonants and the subsections are divided based on manner of articulation.

The consonant inventory of Ulubuya is summarised in (5). There is some controversy in Bantu languages if nasal-consonant sequences should be regarded as a single co-articulated consonant or as a consonant cluster (Kula 1999). Regarding Ulubuya, the argument can be made that they are clusters on the basis of their limited distribution, their complementary distribution with voiceless plosives and the fact that Ulubuya does not exhibit any voiced plosives. Nasal-consonant sequences are discussed separately and in further detail in section 5.2.2.

(5) Ulubuya Consonants

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Palatal	velar
Plosive	p			t			k
Nasal	m			n		ɲ	
Tap or flap				ɾ			
Fricative	β	f		s	ʃ		x
Approximant	w[6]					j	
Lateral approximant				l			
Affricate				ts	tʃ		

The distribution of the consonants is specified below and they are ordered according to their manner of articulation, more specifically being plosive, nasal, liquid, fricative, glide and affricate. A section for nasal-consonant sequences is also included.

6 Voiced labial-velar approximant

5.2.1. **Distribution of plosives**

Ulubuya exhibits a series of voiceless plosives and voiced counterparts are only found in nasal-consonant sequences. The voiceless series also include plosives pronounced with a significant, however non-contrastive, aspiration. No regularities in the distribution were found in the data.

/p/

According to the data, /p/ can be found in both C_1 and C_2 position.

1) C_1 distribution of /p/

a) pá:pa	/pa:pa/	'father' (ncl 1)[7]
b) ipusi	/i-pusi/	'cat' (ncl. 9)
c) xuxu:pisá	/xu-xu:-pisa/	'to urinate' (ncl. 15)
d) lípe:ɾa	/li-pe:ɾa/	'guava fruit' (ncl. 5)

2) C_2 distribution of /p/

a) xú:xupa	/xu:-xupa/	'to beat' (ncl. 15)
b) pá:pa	/pa:pa/	'father' (ncl. 1)

7 ncl.: Noun class. Nouns in Bantu languages are subdivided into noun classes, loosely based on the meaning of the words. This is marked by noun class prefixes and sometimes also includes what is known as a pre-prefix. This is also true for Ulubuya. In the data, whenever a lexeme is segmented into two morphemes, the first morpheme is the noun class prefix and the second is the root (see 1)d)). When segmented into three morphemes, the first morpheme is a pre-prefix, the second is a prefix and the third is the root of the word (see 1)c)).

/t/

The phoneme /t/ is found in C_1 but in the data no instance of /t/ in C_2 was found.

3) C_1 distribution of /t/

a) *ʃítá:nda*	/ʃi-tanda/	'bed construction' (ncl. 7)
b) *xú:tewa*	/xu:-tewa/	'to cook' (ncl. 15)
c) *βu:tí:ní*	/βu:-ti:ni/	'smallness' (ncl. 14)
d) *lí:tú:li*	/li:-tu:li/	'group of people' (ncl. 5)

/k/

/k/ is found in both C_1 and C_2 but its distribution in the data is far more limited in C_2.

4) C_1 distribution of /k/

a) *xú:káni*	/xu:-kani/	'to love, to like, to want' (ncl. 9)
b) *kumuka:ŋgu*	/ku-mu-kangu/	'stick used for cooking' (ncl 3)
c) *ʃi:ke:ɾe*	/ʃi:-ke:ɾe/	'leg' (ncl 7)
d) *úmuko:ko*	/u-mu-ko:ko/	'sister' (ncl 1)
e) *lú:kisu*	/lu:-kisu/	'Lukisu' (language) (ncl 11)

5) C_2 distribution of /k/

a) *í:mbáko*	/í-N-pako/[8]	'hoe' (ncl 9)
b) *βáko:ko*	/βa-ko:ko/	'sisters' (ncl 2)

In conclusion, Ulubuya only exhibits a single series of voiceless plosives. This is in sharp contrast to the general characteristics of many modern day Bantu languages which, just like Proto-Bantu, contain two series of plosives, differentiated through phonation (Maddieson 2003). It is interesting that Ulubuya completely lacks this fairly common feature of many

---

8 N = homorganic nasal. See section 5.2.2. for discussion.

other Bantu languages. Instead, Ulubuya displays a fairly large set of fricatives and affricates, as is discussed in sections (5.2.3.) and (5.2.4.) below.

5.2.2. **Distribution of nasal-consonant sequences**

This section discusses the nasal-consonant sequences found in Ulubuya and if these ought to be analysed as individual segments or consonant clusters. The section also deals with the distribution of these nasal-consonant sequences.

The nasal-consonant sequence (consisting of a nasal and an obstruent consonant, henceforth known as NC) in Bantu languages have generally been analysed either as pre-nasalised consonant or as a consonant cluster. The distribution of NC sequences is presented below (6) since their distribution is relevant for the argumentation that they are clusters.

(6) Distribution of NC sequences in Ulubuya

	Across Morpheme boundary	Stem-initial	Stem-internal
mb	x	-	x
nd	x	-	x
ŋg	x	-	x
ɲʤ	x	-	-
nts	x	-	-

My preliminary analysis of the NC sequences in Ulubuya is that there are two kinds of NC sequences, both clusters and individual segments. The arguments favouring either position seem to be insufficient to justify the analysis that all NC sequences are either clusters or individual segments. Below follows a discussion of the various positions as well as a summary of the discussion.

The NC sequences only appear across morpheme boundaries and stem internally. This lends support to the analysis that they are consonant clusters rather than single segments since the other consonants all occur in the stem-initial position. Furthermore, Ulubuya does not have voiced plosives except in NC sequences and this might be analysed as a NC cluster where an underlying voiceless plosive is realised as voiced due to the effect of the nasal. This analysis is supported by the fact that Ulubuya lacks NC sequences with voiceless plosives.

If one assumes that Bantu languages only have open syllables and that NC sequences are clusters, this would lead to a violation of the sonority sequencing principle (Downing 2005). This could be avoided by abandoning the traditional analysis that Bantu languages only have closed syllables and analysing NC sequences as cluster occurring across syllables. The NC sequence can thus be analysed as CV**N.C**V rather than **NC**V and would not contradict the sonority sequencing principle.

Regarding NC sequences that are distributed across morpheme boundaries, it is common in Bantu languages for them to be analysed as clusters. This is especially true when the noun class prefix for classes nine and ten, consisting of a homorganic nasal, is combined with a stem-initial consonant. When this homorganic nasal is combined with a voiceless plosive it can cause the plosive to become voiced (Hyman 2003). Ulubuya exhibits these attributes and the homorganic nasal can be realised as [m], [n] and [ŋ] as seen in 6)a), 7)a) and 8)a), respectively. The voiced plosive in such sequences has thus been analysed as derived. The analysis that all NC sequences are such clusters would have the benefit of being comparatively economical since one could avoid granting nasal-obstruent sequences independent phonemic status. This would also allow all NC sequences to be analysed in the same manner.

In summation, the arguments for a cluster analysis in Ulubuya are the limited distribution of the NC sequences, the absences of such sequences with voiceless plosives, the absences of independent voiced plosives, being a comparatively economical solution as well as being able to avoid violating the sonority sequencing principle.

The argument can also be made that morpheme-internal NC sequences in Ulubuya ought to be analysed as single segments. This is often how Bantu languages are analysed (Hyman 2003). In Bukusu, Mutonyi (2000) claims that the phonemic status pre-nasalised obstruents can be made on the basis of their shared distribution in intervocalic contexts with the voiceless plosives in Bukusu. Thus, Bukusu is analysed by Mutonyi (2000) as having a series pre-nasalised voiced plosives that is contrasted with the series of voiceless plosives, rather than having two series of plosives, one voiced and one voiceless. Seeing the similarities between Ulubuya and Bukusu, the argument is easily made that a similar analysis ought to be used for Ulubuya.

Furthermore, no language data have been found that would conclusively prove that the obstruents in NC sequences are underlying voiceless plosives. In addition, the argument that a single segment analysis would result in a contradiction of the sonority sequencing principle is not particularly strong since it is a principle that many languages follow rather than a rule that all languages have to adhere to.

Considering the argumentation above, the case for the analysis that all NC sequences ought to be analysed as clusters does not seem to be any stronger than the analysis that Ulubuya also have NC sequences consisting of a single segment. Thus, an attempt at a compromise between the two positions was made. NC sequences with a homorganic nasal that occur across morpheme boundaries have been analysed as clusters with an underlying voiceless plosive since a strong argument can be made for this analysis. As for stem-internal NC sequences, they are presented as a single segment and the obstruents have been presented in the data as underlying voiced plosives.

An accurate and exhaustive analysis of the NC sequences in Ulubuya is beyond the scope of this study and the analysis provided here is merely a preliminary sketch of how the NC sequence might be analysed.

[mb]

6) distribution of [mb]

a) tsí:mbi:tsi	/tsi-N-pi:tsi/	'pigs' (ncl 10)
b) á:mbi	/ambi/	'near'
c) βijé:mbe	/βi-jembe/	'mango trees' (ncl 8)
d) -βwí:mbi	/-βwimbi/	'short'

[nd]

7) distribution of [nd]

a) í:ndé:mo	/i-N-te:mo/	‘snake’ (ncl 9)
b) úmú:ndu	/u-mu-ndu/	‘person’ (ncl 1)
c) úmúsi:nde	/u-mu-sinde/	’boy’ (ncl 1)
d) ʃítá:nda	/ʃi-tanda/	’bed construction’ (ncl 7)

[ŋg]

8) distribution of [ŋg]

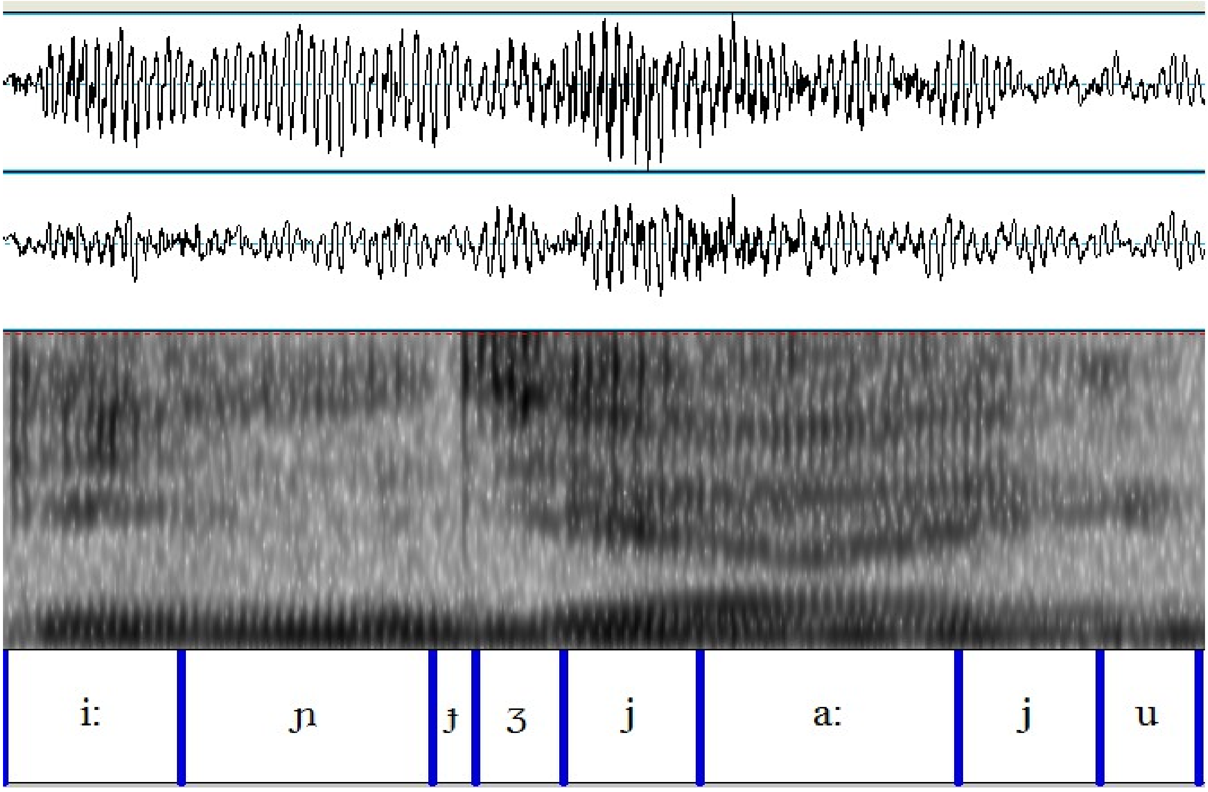
a) í:ŋgá:fu	/i-N-ka:fu/	’cow’ (ncl 9)
b) xú:sá:ŋga:la	/xu:-sanga:la/	’to be happy’ (ncl 15)

[ɲɟʒj]

An interesting phonetic phenomenon was observed when the homorganic nasal noun class prefix preceded a stem initial /j/. The place of articulation of the homorganic nasal was assimilated to the following underlying palatal glide /j/ and realised as /ɲ/. Then a hardening of the glide followed, resulting in an affricate [ɟʒ], followed by a glide /j/, before reaching the first vowel. Evidence for the nature of the underlying form is found in 9)c), where the same stem is combined with another prefix.

9) /ɲɟʒj/

a) Annotated spectrogram of *[í:ɲɟʒjá:ju]*



b) *í:ɲɟʒjá:ju* */i-N-jaju/* 'wild cat' (ncl. 9)

c) *tʃimijá:ju* */tʃi-mi-jaju/* 'wild cats' (ncl. 4)

5.2.3. **Distribution of fricatives and spirantisation**

This section deals with the distribution of fricatives in Ulubuya. It also discusses the voiced sibilant [z] found in the data and argues against its phonemic status. A subsection deals with the historical process of spirantisation.

/β/

There is some free variation regarding the phoneme /β/ in the data. In a word initial position, it can also be realised as a voiced bilabial stop rather than as a fricative, although the stop is not as forceful as the stop found in the NC sequence [mb].

10) C_1 distribution of /β/

a) kámáβa:ɾi	/ka-ma-βa:ɾi/	’stones’ (ncl 6)
b) lúβu:ja	/lu-βu:ja/	’Ulubuya’ (language) (ncl 11)
c) kúmúβiɾi	/ku-mu-βiɾi/	’body’ (ncl 3)

11) C_2 distribution of /β/

a) tsí:ŋgóβe	/tsi-N-koβe/	’monkeys’ (ncl 10)
b) xú:xwiβa	/xú:-xu-iβa/	‘to steal’ (ncl. 15)

/f/

12) distribution of /f/

a) í:ntsófu	/i-N-tsofu/	‘elephant’ (ncl 9)
b) í:ŋgá:fu	/í-N-ka:fu/	’cow’ (ncl 9)
c) βúnífu	/βu-nifu/	‘dampness' (ncl 14)

/x/

13) C_1 distribution of /x/

a) kúmúxu:nu	/ku-mu-xu:nu/	‘arm, hand’ (ncl 3)
b) ʃíxo:βa	/ʃi-xo:βa/	‘skin’ (ncl 7)
c) úmúxa:na	/u-mu-xa:na/	‘girl’ (ncl 1)

14) C_2 distribution of /x/

a) lúβó:xa	/lu-βo:xa/	‘algae’ (ncl 11)
b) xú:lexa	/xu:-lexa/	’to leave’ (ncl. 15)

/s/

15) C_1 distribution of /s/

a) βú:síma	/βu:-sima/	(name of a dish) (ncl 14)
b) βú:sú:ŋgu	/βu:-sungu/	’anger’ (ncl 14)

16) C_2 distribution of /s/

a) xú:kwisa	/xu:-kwisa/	‘to drop’ (ncl. 15)
b) kúmúje:se	/ku-mu-je:se/	‘the moon’ (ncl 3)
c) tsípusi	/tsi-pusi/	‘cats’ (ncl. 10)

[z]

There was only one example of [z] found in the data and this example was not found inside a stem. Furthermore, since Ulubuya neither has plosives or fricatives that are distinguished merely by phonation, a strong argument can be made that [z] is an allophonic variation of /s/ rather than an independent phoneme. According to the native speaker however, the stem of the word for anger (*βu:sú:ŋgu,* see 15)b)) and the stem of the word for white person (*úmúzu:ŋgu*, see 17)a)) constitute a minimal pair in Ulubuya. This might be a foreign loan word from the Swahili word *muzuŋgu* (white person) and therefore it is not analysed as a separate phoneme in Ulubuya. [9]

17) C_1 distribution of [z]

a) úmúzu:ŋgu	/u-mu-zungu/	‘white person’ (ncl 1)

/ʃ/

18) C_1 distribution of /ʃ/

a) ʃí:na	/ʃi:na/	‘what’
b) xú:ʃína	/xu:ʃina/	‘to dance’ (ncl. 15)
c) xú:ʃá:ma	/xu:ʃa:ma/	‘to defecate’ (ncl. 15)

9 I would like to thank Mirjam Möller for pointing this out and advising me on the matter.

19) C_2 distribution of /ʃ/

a) lí:koʃe	/li:-koʃe/	'ash' (ncl. 5)
b) lwáʃĭ	/lwaʃi/	'why'

It is interesting that apart from a voiced bilabial fricative, Ulubuya does not have any voiced fricatives. Additionally, Ulubuya does not contain any fricatives that are distinguished merely by phonation since it lacks a voiceless bilabial fricative. A similar pattern is also found in the set of plosives in Ulubuya, which also lacks voiced counterparts.

Spirantisation

In this section I discuss the historical process of spirantisation within Bantu languages as well as provide examples that attempt to show the Ulubuya has undergone a historical process of spirantisation.

Spirantisation[10], the process in which a plosive turns into a fricative (Trask 2012), is attested among many of the world's languages, including many Bantu languages. A brief presentation of the key features of spirantisation in Bantu languages will follow in order to further analyse the data regarding Ulubuya and contrast it to reconstructions of Proto-Bantu.

Among Bantu languages, spirantisation has occurred with plosives preceding high vowels in Proto-Bantu (see section 4.2). Although the reconstruction of Proto-Bantu has seven vowels, many Bantu languages have undergone a change into a five vowel system due to the loss of the extra high vowels (Hyman 2003). Almost all languages that have undergone the change into a system of five vowels show signs of spirantisation (Janson 2007). This primarily concerns the first consonant of a word stem and is summarised in 20).

10 Using the term 'spirantisation' in this context is somewhat problematic, as also discussed by Janson (2007). Traditionally the term 'spirantisation' only covers the transition from a plosive (often via an affricate) to a fricative whereas the data used by Janson (2007), among others, also include changes in place of articulation. The wider term 'lenition' might be more appropriate. In this study however, according to the praxis already established by many bantuists, (e.g. Janson 2007, Maddieson 2003) the term 'spirantisation' is used.

20) Spirantisation changes in Bantu Languages

	Before i	before u	
p, b >	f, v (or, s, z)	f, v	
t, l[11] >	s, z	f, v (or, s, z)	
k, g>	s, z	f, v	(Schadeberg 1994)

There are, however, intermediate steps in the process of spirantisation that occur before the process is completed. The consonants found in 20) only show the end results. According to Janson (2007) a common intermediate step is the change from plosive to affricative, before finally resulting in a fricative. Among others, *ti > tsi > si*, *bi > bzi > zi* and *tu > tʃu >ʃu > fu* are given as examples (Janson 2007). Present day Bantu Languages are to be found all along this spirantisation continuum.

According to Maddieson (2003), typologically common fricatives are comparatively rare in Bantu languages, although he does not specify which fricatives these are. This is not the case for Ulubuya which, at least in comparison to other Bantu languages, has a rather large set of fricatives. Considering spirantisation in the Bantu language family, this might be related to the fact that Ulubuya only has five vowels as well as only a single series of plosives. The comparatively large set of fricatives might be in response to this. Since almost all Bantu languages that have undergone spirantisation also have had its set of vowels reduced from seven to five, this analysis of spirantisation Ulubuya seems to fit in well with general Bantu phonology.

This analysis is further supported by the lexical reconstruction of Proto-Bantu. A reconstruction of the word for 'elephant' in Proto-Bantu is **jògù* (Bantu lexical reconstructions 3 2014) where in Ulubuya the stem of the word is *–tsófu*. It would seem that there has been a process of spirantisation where the **g* found in Proto-Bantu has changed into /f/ in modern day Ulubuya.

Further evidence of spirantisation is found in the stem of the word *-síma* (the name of a dish) in Ulubuya . In the lexical reconstruction for Proto-Bantu, the word is **kímà* (Bantu lexical

11 Here [l] is used instead of [d] as seen in section 4.2. There is, as also mentioned in section 4.2, some disagreement regarding whether or not this consonant should be considered an alveolar plosive or sonorant.

reconstructions 3 2014) is given as the name for the same dish. Here the reconstructed **/k/* is proposed to have changed into /s/ in Ulubuya.

Another process in which a plosive has transformed into a fricative is observed in the reconstruction of the word for anger, which is **cúngú*, (Bantu lexical reconstructions 3 2014) and the equivalent in Ulubuya, *-súngu*. This seems to be in accordance with Schadeberg's (1994) discussion of spirantisation within the Bantu language family. In the data however, only a few such examples have been found and they are not enough on their own to constitute evidence sufficient to draw any conclusions on. More data and further studies are needed in order to properly establish a theory regarding this. However, several instances of words have been found, whose diachronic change is in line with the principles laid down by Janson (2007) and the analysis was made that Ulubuya has undergone a process of spirantisation.

5.2.4. **Distribution of affricates**

In this section I write about the distribution of affricates and discuss the possibility that the affricates in Ulubuya are the result of spirantisation. Ulubuya has two affricates, /ts/ and /tʃ/, both of which are voiceless. A voiced affricate was found in the data, but that was the result of a prefix ending in with a homorganic nasal combining with a stem-initial /j/ (see section (5.2.2) for a discussion and 9) for examples).

An interesting feature is that even when the affricate /ts/ occurred after the homorganic nasal it remained unvoiced, in contrast to what happens with unvoiced plosives. For examples, see /ts/ in 21)a) and /k/ (which becomes voiced) in 8)a) where both are examples contain the noun class 9/10 prefix.

/ts/

21) C_1 distribution of /ts/

a) tsí:ntsófu	/tsi-N-tsofu/	'elephants' (ncl. 10)
b) xú:tsú:ŋga	/xu:-tsunga/	'to pound' (ncl 15)
c) kúmútsu:ŋgiɾu	/ku-mu-tsungiɾu/	'stick for pounding'(ncl 3)

22) C_2 distribution of /ts/

a) *βaxó:tsa*	/βa-xo:tsa/	'uncles' (maternal) (ncl 2)
b) *í:mbitsi*	/i-N-pitsi/	'pig' (ncl 9)

/tʃ/

There are relatively few examples of the phoneme /tʃ/ belonging to the stem of a word in the data. The affricate is part of the pre-prefix of the noun class 4 prefix, meaning that several instances of /tʃ/ are still found throughout the data.

23) distribution of /tʃ/

a) *lí:tʃi*	/li:-tʃi/	'egg'	(ncl 5)
b) *kamatʃi*	/ka-ma-tʃi/	'eggs'	(ncl 6)
c) *tʃimiβiɾi*	/tʃi-mi-βiɾi/	'bodies'	(ncl 4)
d) *tʃimíjé:mbe*	/tʃi-mi-je:mbe/	'mango fruits'	(ncl 4)

According Janson (2007), the spirantisation of plosives into fricatives is a gradual process, where affricates are included as intermediate steps in the process. The stem of the word in Ulubuya for 'to pound' is *-tsú:ŋga* and the word reconstructed for Proto-Bantu Zone J is **tóang*. The affricate /ts/ in Ulubuya might be the result of a spiratisation process of the Proto-Bantu **/t/.* This is only a single word and as mentioned above, further data and studies are needed in order to draw any conclusions regarding this.

5.2.5. Distribution of nasals

In this section I write about the distribution of nasals in Ulubuya and provide arguments for the phonemic status of the palatal nasal /ɲ/. Ulubuya has a set of three nasals; bilabial, alveolar and palatal. Nasals are found throughout both C_1 and C_2.

/m/

24) C_1 distribution of /m/

a) xú:maɲa	/xu:maɲa/	'to know' (ncl. 15)
b) ʃímóli	/ʃi-moli/	'flower' (ncl. 7)

25) C_2 distribution of /m/

a) xú:lima	/xu:-lima/	'to cultivate' (ncl. 15)
b) xúxwo:ma	/xu-xu-oma/	'to dry' (ncl. 15)

/n/

26) C_1 distribution of /n/

a) βu-nifù	/βu-nifu/	'dampness' (ncl. 14)
b) ná:nu	/na:nu/	'who'

27) C_2 distribution of /n/

a) úmwa:na	/u-mu-ana/	'child' (ncl. 1)
b) βu:tí:ní	/βu:-ti:ni/	'smallness' (ncl. 14)
c) úxú:iβe:na	/u-xu:-iβe:na/	'to bleed' (ncl. 15)
d) kúmúxu:nu	/ku-mu-xu:nu/	'arm, hand' (ncl. 3)

/ɲ/

28) distribution of /ɲ/

a) βu:ɲasi	/βu:-ɲasi/	'grass' (ncl. 14)
b) í:ɲama	/í-ɲama/	'meat' (ncl. 9)
c) í:ɲá:ŋga	/i:-ɲanga/	'sun' (ncl. 9)
d) xú:maɲa	/xu:maɲa/	'to know' (ncl. 15)

It was difficult to determine whether the words in 28) ought to be analysed as containing [ɲj] or/ɲ/ based on perception alone. In order to lend further support to the analysis that /ɲ/ is an independent phoneme in Ulubuya, the stems of some of the words given in 28) are compared below with their reconstructed equivalents for Proto-Bantu in 29), all of which contain the phoneme /ɲ/. These words probably have a historical connection, thus supporting analysis that /ɲ/ is a phoneme in Ulubuya.

Additionally, the nasal-glide sequence that was found in the data resulted in a hardening of the glide, causing it to be realised as an affricate. This lends further support to analysis that the words given in 28) does not contain a sequence consistiting of [ɲj] but rather the phoneme /ɲ/.

29) /ɲ/ in Ulubuya and reconstructed Proto-Bantu

a) -ɲasi	**nyàkí*[12]	'grass'
b) -ɲama	**nyàmà*	‘meat’
c) -maɲa	**màny*	‘to know’

(Bantu lexical reconstructions 3 2014)

Many Bantu languages contain a full set of nasals corresponding to place of articulation where fricatives and plosives are to be found (Maddieson 2003). By this standard, Ulubuya should also include a velar nasal since Ulubuya has both a velar fricative and a velar plosive. However, in Ulubuya the velar nasal [ŋ] is only found in NC sequences and has thus been analysed as the result of a homorganic nasal combined with a velar plosive rather than as an independent phoneme. In contrast, Bukusu contains a full set of nasals, including the velar nasal missing in Ulubuya (Mutonyi 2000).

5.2.6. **Distribution of glides**

This section discusses derived and underlying glides, the two kinds of glides found in Ulubuya. This section gives examples of derived glides and also includes the distribution of underlying glides. The data is also compared to glides in Bukusu.

12 /ny/ represents /ɲ/ in the reconstructions for Proto-Bantu. (Bantu lexical reconstructions 3 2014)

Mutonyi (2000) claims that Bukusu has at least three kinds of glides; underlying, derived and epenthetic. Derived glides are formed when the high vowels /i/ and /u/ glide before other vowels. /i/ is realised as /y/[13] and /u/ is realised as /w/ when they occur before other vowels. Underlying glides however, occur both stem-initially and stem-internally and are not the result of glide formations (Mutonyi 2000).

Based on the data, Ulubuya exhibits similar traits with both derived and underlying glides. A derived glide can be found in the word *úmwa:na* (child), where the underlying prefix /*mu*-/ is realised as [*mw*-]. This does not occur when the prefix precedes a consonant initial stem, as seen in *úmúlimi* (farmer). Glide formation in Ulubuya seem to cause a lengthening of the following vowel. A similar feature is observed in Bukusu (Mutonyi 2000).

In 30) examples both derived /w/ and /j/ are given in addition to examples of the same roots and prefixes where glides formation does not occur in order to provide evidence for their underlying forms.

30) derived glides in Ulubuya

a) úmwa:na	/u-mu-ana/	'child' (ncl. 1)
b) áβa:na	/a-βa-ana/	'children' (ncl. 2)
c) úmúlimi	/u-mu-limi/	'farmer' (ncl. 1)
d) liljé:no	/li-li-eno/	'tooth' (ncl. 5)
e) kámé:no	/ka-ma-eno/	'teeth (ncl. 6)
f) li:βá:le	/li:-βa:le/	'stone' (ncl. 5)

Ulubuya has both /j/ and /w/ as underlying glides. Both glides appear stem-internally and /j/ also occurs stem-initially and can therefore, on the basis of distribution, be analysed as underlying. Ergo, they are treated as independent phonemes in Ulubuya.

13 /y/ is used by Mutonyi (2000) to represent the IPA symbol /j/ and when referring to Bukusu in this paper I have adopted the praxis used by Mutonyi (2000).

/j/

31) C_1 distribution of /j/

a) *kúmújá:ju* /ku-mu-ja:ju/ ’wild cat’ (ncl. 3)

b) *í:ɲʒɟjá:ju* /i-N-ja:ju/ ‘wild cat’ (ncl. 9)

c) *kúmúje:se* /ku-mu-je:se/ ‘the moon’ (ncl 3)

d) *βá:jina* /βá:jina/ ’where’ (specific location)

32) C_2 distribution of /j/

a) *xú:ɾija* /xu:ɾija/ ‘to fear’ (ncl. 15)

b) *βu:lá:jí* /βu:la:ji/ ‘goodness’ (ncl. 14)

c) *βu:lé:jí* /βu:le:ji/ ‘tallness’ (ncl. 14)

d) *luβú:ja* /lu-βu:ja/ ’Ulubuya’ (language) (ncl. 11)

/w/

33) distribution of /w/

a) *xú:tewa* /xu:tewa/ ’to cook’ (ncl. 15)

b) *úmúɾówa* /u-mu-ɾowa/ ‘younger sibling’ (ncl. 1)

The glides /w/ and /j/ are a common feature within Bantu languages and often alternate with the high vowels /u/ and /i/, respectively (Maddieson 2003). This is, as seen above, also true for Ulubuya which include the same two glides found throughout many other Bantu languages and also have the same derivation of glides from the high vowels /u/ and /i/. Regarding glides, Ulubuya fits well into the generalisations laid down by Maddieson (2003) on Bantu languages.

5.2.7. **Distribution of liquids**

In this section I write about the distribution of liquids in Ulubuya and discuss whether or not there is any alternation between the two as is the case in Bukusu.

Bukusu has two liquids, /r/ and /l/, and Mutonyi (2000) claims that some speakers alternate freely between them if the underlying phoneme is /r/. In Ulubuya, just as in Bukusu, the two liquids are separate phonemes and this analysis is supported by the minimal pair *xú:ɾija* (to fear) and *xú:lija* (to eat') (see examples 34)a) and 36)a)). No other minimal pairs were found in the data.

During the course of the interviews, at times the native speaker seemed to alternate between the two liquids and at different recordings provided different pronunciations of the same word. Hence, during one of the sessions I brought up previously elicited words containing liquids and attempted to alternate between the liquids. I was corrected by the native speaker, telling me that the attempted alternations were not allowed. This lends further support to the analysis that they separate phonemes. However, based Mutonyi's (2000) description of the liquids in Bukusu and the uncertainty of language data collected on Ulubuya, there might be some alternation between the two liquids in Ulubuya. This is a topic in need of further research.

In the Bantu Language family, most commonly there is only one kind of liquid to be found in each language (Maddieson 2003). This is, as seen below, not the case for Ulubuya . It is interesting then that in both in Ulubuya and Bukusu, both of which contain more than one liquid, these two phonemes seem to allow alternation to some extent.

/ɾ/

34) C_1 distribution of /ɾ/

a) xú:ɾija	/xu:ɾija/	'to fear' (ncl. 15)
b) kámáɾoɾo	/ka-ma-ɾoɾo/	'dreams' (ncl. 6)
c) liɾóβa	/li-ɾoβa/	'ground' (ncl. 5)

35) C2 distribution of /ɾ/

a) *kámápe:ɾa* /ka-ma-pe:ɾa/ ‘guavas’ (ncl. 6)

b) *kúmúliɾu* /ku-mu-liɾu/ ’fire' (ncl. 3)

c) *tʃímíβiɾi* /tʃi-mi-βiɾi/ ’bodies’ (ncl 4)

d) *ʃíɾóɾo* /ʃi-ɾoɾo/ ‘dream’ (ncl. 7)

/l/

36) C1 distribution of /l/

a) *xú:lija* /xu:lija/ ‘to eat’ (ncl. 15)

b) *βu:lá:jí* /βu:la:ji/ ‘goodness’ (ncl. 14)

c) *βu:lé:jí* /βu:le:ji/ ‘tallness’ (ncl. 14)

37) C2 distribution of /l/

a) *xú:βula* /xu:-βula/ ‘to defeat’ (ncl. 15)

b) *βímóli* /βi-moli/ ‘flowers’ (ncl. 8)

c) *xú:lila* /xu:lila/ ‘to cry’ (ncl. 15)

5.3. Vowels

In this section I write about the vowel phonemes found in Ulubuya. The section also discusses the environments in which vowel lengthening occur and includes a vowel chart.

The vowel system of a Bantu language typically consists of the typologically common five vowel system /i, e, a, o, u/ (Andersson 2001). This is, however, somewhat of a simplification, as pointed out by Maddieson (2003), where the phonetic realisation of the phonemes /e/ and /o/ display very different formant values in different languages but are still identified with the IPA symbol /e/ and /o/, respectively. Another common set of vowels among Bantu languages is a set of seven vowels (Maddieson 2003). Languages with nine and more vowels are to be found among Bantu languages but they are, however, not as common as the five and seven vowel systems (Hyman 2003).

Ulubuya has a standard set of five vowels (/i/, /e/, /a/ /o/ and /u/). The data includes both long and short vowels. Long vowels preceding NC sequences, following derived glides and in prefix hapolology have been analysed as underlying short vowels. Vowel lengthening in such environments is a common feature in Bantu languages and is also present in Bukusu (Mutonyi 2000). In Ulubuya, vowels found in such environments were all long. In Bukusu short vowels all have long counterparts but are not treated as separate phonemes by Mutonyi (2000).

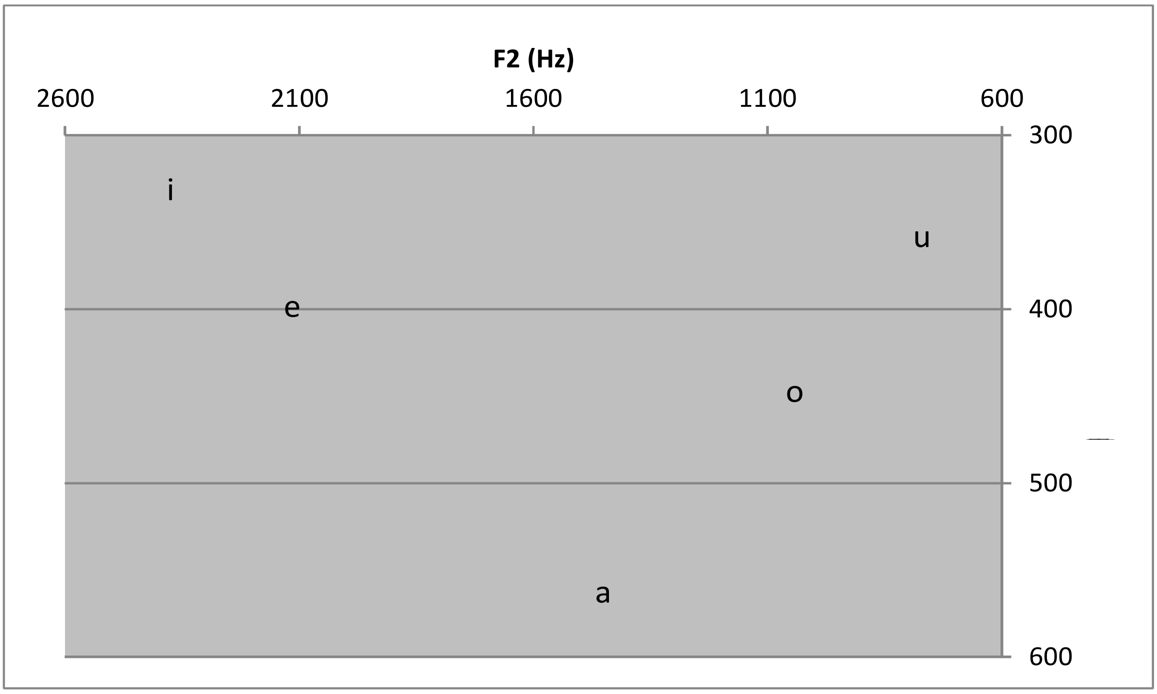
It is beyond the scope of this study to provide an accurate description of the phonetic function of vowel length in Ulubuya. However, that fact that long vowels are found in the data in environments where length is not predictable provide a good argument for the interpretation that vowel length is a contrastive in Ulubuya. 38) includes examples of each long vowel in Ulubuya in environments where length is not predictable.

38) Long vowels

a) xú:βula	/xu:-βula/	‘to defeat’ (ncl. 15)
b) βu:lá:jí	*/βu:la:ji/*	‘goodness’ (ncl. 14)
c) kámápe:ɾa	/ka-ma-pe:ɾa/	‘guavas’ (ncl. 6)
d) βu:tí:ní	/βu:-ti:ni/	’smallness’ (ncl. 14)
e) úmuko:ko	/u-mu-ko:ko/	‘sister’ (ncl 1)

Figure (7) below is a vowel chart for Ulubuya. The data was collected by measuring the vowel formants of words spoken both in isolation and within phrases. The program Praat was used to measure the vowel formants. Then the vowel formant means were calculated and presented in the chart below. The chart displays the F1 and F2 value of the vowels in Hz in order to display their value in relation to each other and their locations within the vowel space.

(7) Ulubuya vowel formant means[14].



5.4. Syllable structure

Here I write about the different syllables found in Ulubuya. I also discuss the possibility of closed syllables in Ulubuya.

The syllable structure of many modern Bantu languages closely resembles that of Proto-Bantu (Hyman 2003). Even though the syllable structure of several Bantu languages has undergone changes, modern day Bantu languages still adhere to the comparatively small variation of syllables found in the reconstructed Proto-Bantu. This feature is reinforced by the agglutinating, predominantly prefixing, nature of Bantu morphology. There are examples of languages that have lost the final vowel, resulting in the occurrence of closed syllables, although the open syllable feature is still very common. Generally due to the loss of vowels or consonants, there are also non-final closed syllables to be found in some modern day Bantu languages. In some cases this is the result of borrowings from other language families (Hyman 2003).

14 Each point represent the average measurement of at lest 21 tokens spoken by a male speaker.

Most syllables in Ulubuya are open. The claim that Ulubuya has closed syllables is dependent upon the analysis that NC sequences are clusters and not pre-nasalised consonants. Kula (1999) proposes that Bantu languages can have the closed CVN structure. Ndengleko is an example of Bantu language analysed as having closed syllables (Ström 2013). Furthermore, Downing (2005) also argues that Bantu languages can have the closed CVN structure.

As discussed in section 5.2.2., it is difficult to build up a strong argument for the analysis that Ulubuya has morpheme-internal NC clusters and closed syllables. It is beyond the scope of this study to determine whether or not Ulubuya has closed syllables and they are included in (8) in brackets for the sake of completion.

(8) Syllables in Ulubuya

a) Open Syllables	V, CV, CVV, NCV, GV, CGV[15]
b) Closed Syllables	(VN),(CVN), (GVN)

The consonant in CV(V) can be a nasal. The syllables in Ulubuya can have complex onsets but, under the assumption that Ulubuya has closed syllables, does not exhibit any complex codas. Consonant cluster can occur across morpheme boundaries.

5.5. Tone

In this section I summarise what was discovered about tone in Ulubuya. I primarily relate the distribution of high and low tones within a stem and this is compared to the tonal structures of Proto-Bantu and other modern day Bantu languages.

According to Katamba (1989) many African languages have grammatical, rather than lexical tone. The two are not necessarily mutually exclusive however, and Luganda is an example of a language with both grammatical and lexical tone (Katamba 1989). Kissebirth & Odden (2003) lend support to this view and beyond lexical tones also claim that tone as a distinguishing feature within many Bantu languages is related to syntactic and morphological processes. According to Mutonyi (2000), Bukusu has retained the High/Low lexical contrast from Proto-Bantu, at least for its verbs. Noun phrases also distinguish between high and low tone.

15 V:Vowel, C:Consonant, G:Glide, N:Nasal

Many Bantu languages have tone as a distinguishing feature, but this is not true for all languages and Swahili, among others, is a non-tonal language (Kisseberth & Odden 2003). Usually high and low (sometimes described as toneless) tone are distinguished but more complex systems are also found, including a super high tone and a super low tone, in addition to the distinction between high and low. As discussed above, the canonical stem is disyllabic so for the reconstruction of Proto-Bantu, four patterns have be reconstructed, namely High-High, High-Low, Low-Low and Low-High (Kisseberth & Odden 2003).

Based on the data, Ulubuya distinguishes between high and low tone in a similar manner to many other Bantu languages. It is not currently clear whether or not the low to should be regarded as truly low tone or as a toneless counterpart to the high tone.

Ulubuya shows pairs of syllables which exhibit all combinations that can be found based on the high-low distinction, being High-High, High-Low, Low-Low and Low-High. Examples of each combination are presented in 38). The exact function of the tones in Ulubuya is still unclear and is beyond the scope of this study.

39) Possible combinations of tones within a stem

a) βu:tí:ní	/βu:-ti:ni/	‘smallness’ (ncl. 14)	(High-High)
b) pá:pa	/pa:pa/	‘father’ (ncl 1)	(High-Low)
c) βu:ɲasi	/βu:-ɲasi/	’grass’ (ncl. 14)	(Low-Low)
d) xuxu:pisá	/xu-xu:-pisa/	‘to urinate’ (ncl. 15)	(Low-High)

6. A typological perspective

In this section I take a typological perspective to analyse and summarise the phonemes found in Ulubuya. The phoneme inventory is compared both to common features of Bantu languages as well as general phonological typology.

The phoneme inventory of Ulubuya exhibits, from a typological perspective, some interesting phonological features. Considering the size of its consonant inventory, it is interesting that Ulubuya lacks voiced plosives since the distinction between voiced and unvoiced plosives is common among languages with a larger set of consonants (Lindblad 2010). As mentioned above, Ulubuya also has a fairly large set of fricatives, which is interesting in light of

Maddieson's claim (2003) that typologically common types of fricatives are comparatively rare among Bantu languages. The correlation between that small set of plosives and larger set of fricatives and affricates is an interesting one. Furthermore, the general lack of symmetry between voice and voiceless plosives, fricatives and affricates is also interesting since Lindblad (2010) claims that symmetry is a strong tendency among the consonant systems in many of the world's languages. It is interesting that Ulubuya does not have any consonants where phonation is the only distinguishing feature.

Considering the nasals in Ulubuya, the inclusion and distinction between /n/ and /m/ is a common feature throughout the world's languages. The palatal /ɲ/ is less common (Lindblad 2010). However, considering the claim made by Maddieson (2003) that Bantu languages generally have nasals correlating to the places of articulation where plosives and fricatives are found, the set of nasals in Ulubuya is smaller than expected. Ulubuya has both the velar fricative /x/ and the velar plosive /k/ but not the velar nasal /ŋ/. The velar nasal /ŋ/ is, from a typological perspective, a less common feature (Andersson 2001).

The most common set of vowels found in the world's languages is a set of five vowels, consisting of the following /i, e, a, o, u/ and this set often form the base more complex systems (Engstrand 2004). This is also the most common set of vowels found among Bantu languages (Maddieson 2003). The vowel system in Ulubuya is a very common system from a typological perspective.

7. Conclusions and further research

In this section I provide a brief summary of the study, its results and conclusions. I also discuss areas in Ulubuya that still need to be further researched.

This is a study of Ulubuya, a variety within the Masaba dialect continuum and the data which this study is based on was gathered during interviews with a native speaker. At the time of this study the variant was still undocumented and the aim of this study was to make a contribution to our knowledge of the aforementioned dialect continuum as well as to help preserve the variety known as Ulubuya.

Ulubuya has many phonological properties that are typical for Bantu languages but also displays some less common attributes, such as its comparatively large set of fricatives and affricates as well as lacking a distinction between voiced and voiceless plosives. Nasals, the vowel set and syllable structure share many common features found in other Bantu languages.

At the time of this study, Ulubuya was an undocumented variation within the Masaba dialect continuum and several areas of the language variety remain undocumented. From a phonological standpoint, further research into the function of tone is needed in order to provide a comprehensive and accurate description. The function of vowel length is also in need of further research. Allophonic variations also need further investigation, in particular the allophonic variations of /β/ as well as the interchangeability between the liquids /ɾ/ and /l/. In order to provide a comprehensive analysis and description of Ulubuya, areas beyond the scope of phonology also need to be researched, such as the grammar and the lexicography of the language.

8. **References**

Andersson, L.-G. (2001). *Språktypologi och språksläktskap* (4. uppl.). Stockholm: Liber.

Austen, C. L. (1974). *Aspects of Bukusu syntax and phonology*. Indiana University.

Bantu lexical reconstructions 3. (2005). Retrieved 2014-05-09 from http://www.africamuseum.be/collections/browsecollections/humansciences/blr

Blois, K. F. (1975). *Bukusu generative phonology and aspects of Bantu structure*. Koninklijk Museum voor Midden-Afrika.

Bowern, C. (2008). *Linguistic fieldwork: a practical guide*. New York: Palgrave Macmillan.

Chelliah, S. L. (2010). *Handbook of descriptive linguistic fieldwork*. Dordrecht ; New York: Springer.

Downing, L. J. (2005). On the ambiguous segmental status of nasals in homorganic NC sequences. *The internal organization of phonological segments.* Berlin & New York: Mouton de Gruyter, 183-216.

Ethnologue: Languages of the World, Seventeenth edition. (2014). *Masaaba.* Retrieved 2014-05-11 from http://www.ethnologue.com/language/myx

Engstrand, O. (2004). *Fonetikens grunder.* Studentlitteratur.

Glottolog 2.2. (2013). *Masaaba*. Retrieved 2014-05-11 from http://glottolog.org/resource/languoid/id/masa1299

Guthrie, M. (1967). *Comparative Bantu: an introduction to the comparative linguistics and prehistory of the Bantu languages*. London: Gregg.

Hyman, L. M. (2003). Segmental phonology. *Nurse D. and Philippson G. (eds.). The Bantu Languages,* 42–58.

Janson, T. (2007). Bantu spirantisation as an areal change. *Africana Linguistica*, *13*, 79–116.

Katamba, F. (1989). *An introduction to phonology* (Vol. 4). London: Longman.

Kisseberth, C., & Odden, D. (2003). Tone. *Nurse D. and Philippson G. (eds.). The Bantu Languages,* 59-70.

Kula, N. C. (1999). On the representation of NC clusters in Bemba. *Linguistics in the Netherlands*, *16*(1), 135–148.

Lindblad, P (2010). *Fonetikens grunder.* Lunds Universitet.

Maddieson, I. (2003). The sounds of the Bantu languages. *Nurse D. and Philippson G. (eds.). The Bantu Languages,* 15–41.

Maho, J. (2003). A classification of the Bantu languages: an update of Guthrie's referential system. *Nurse D. and Philippson G. (eds.). The Bantu Languages,* 639–651.

Meeussen, A. E. (1967). Bantu grammatical reconstructions. *Africana linguistica*, *3*, 79-121.

Mutonyi, N. (2000). Aspects of Bukusu morphology and phonology. Ohio State University.

Schadeberg, T. C. (1994). Spirantization and the 7-to-5 vowel merger in Bantu. *Belgian Journal of Linguistics*, *9*(1), 73–84.

Schadeberg, T. C. (2003). Historical linguistics. *Nurse D. and Philippson G. (eds.). The Bantu Languages,* 143–163.

Siertsema, B. 1968. *Pros and cons of macro-phonemes in new orthographies: Masaba spelling problems*. *Lingua* 21. 429-442.

Siertsema, B. 1981. *Masaba Word List: English - Masaba, Masaba - English*. Royale de l'Afrique Centrale: Archives d'Anthropologie, 28. Belgique: Tervuren.

Ström, E.-M. (2013). The Ndengeleko Language of Tanzania.

Trask, R. L. (2012). *A dictionary of phonetics and phonology*. Routledge.