The synchronic phonology and nominal morphology of Sagala (Bantu G39, Tanzania): A preliminary study

Flore Bollaert
Student number: 01406829

Supervisor: Prof. Dr. Koen Bostoen

Acknowledgments

Contents

List of tables

List of figures

Abbreviations

5V	five vowel system
7V	seven vowel system
adj.	adjective
adv.	adverb
ANS	Agent Noun Spirantisation
C	consonant
G	glide
intr.	intransitive
IPA	International Phonetic Alphabet
N	nasal
n.	noun
NCS	noun class singular
NCP	noun class plural
NP	nominal class prefix
num.	number
PB	Proto-Bantu
PFV	perfective
pl.	plural
pron.	pronoun
PrPr	pre-prefix
PST	past
SM	subject marker
tr.	transitive
V	vowel
v.	verb

1. Introduction

The focus of this dissertation is twofold. On the one hand, it provides a preliminary description of the synchronic phonology of the language in question, Sagala[1]. In doing so, it provides a discussion of the phonemes, syllable structure, prosody, and certain morpho-phonological rules. On the other hand, this dissertation also presents a first description of Sagala's noun class system.

In the next subsections, Sagala will be situated both geographically and in relation to other languages spoken in its neighbourhood. I also offer a state-of-the-art of existing research on Sagala and explain the goals of this study. Then I elaborate on the course of my research. As a last part of this introductory section, I account for the orthography adopted in this dissertation.

1.1.   *Situating Sagala*

Sagala is a language spoken by an ethno-linguistic group that lives in central-eastern Tanzania, more specifically in the Morogoro region, shown in Figure 1. The regions in which native Sagala speakers predominantly live are heavily mountainous (Gonzales 2002: 9). In a table in *Researching and Documenting the Languages of Tanzania* by Muzale and Rugemalira (2008: 79-80), which shows what they consider to be the full list of 156 languages spoken in Tanzania with their number of speakers, Sagala was shown to have 106,331 speakers, thus ranking 63rd out of 156. It should be noted that the language name in their table is 'Sagara' instead of 'Sagala'. As I explain in the section on phonology, [l] and [r] are interchangeable in Sagala.

---

[1] The language name is used without nominal prefix. Even though the use of a prefix is grammatically obligatory in any specific Bantu language, this text is written in English. It is not necessary to adopt foreign inflectional paradigms when writing English prose (cf. Maho's 2009 New Updated Guthrie List).

Map 7: The Ruvu Peoples and Their Neighbors, c. 1700 - 1900 A.D.
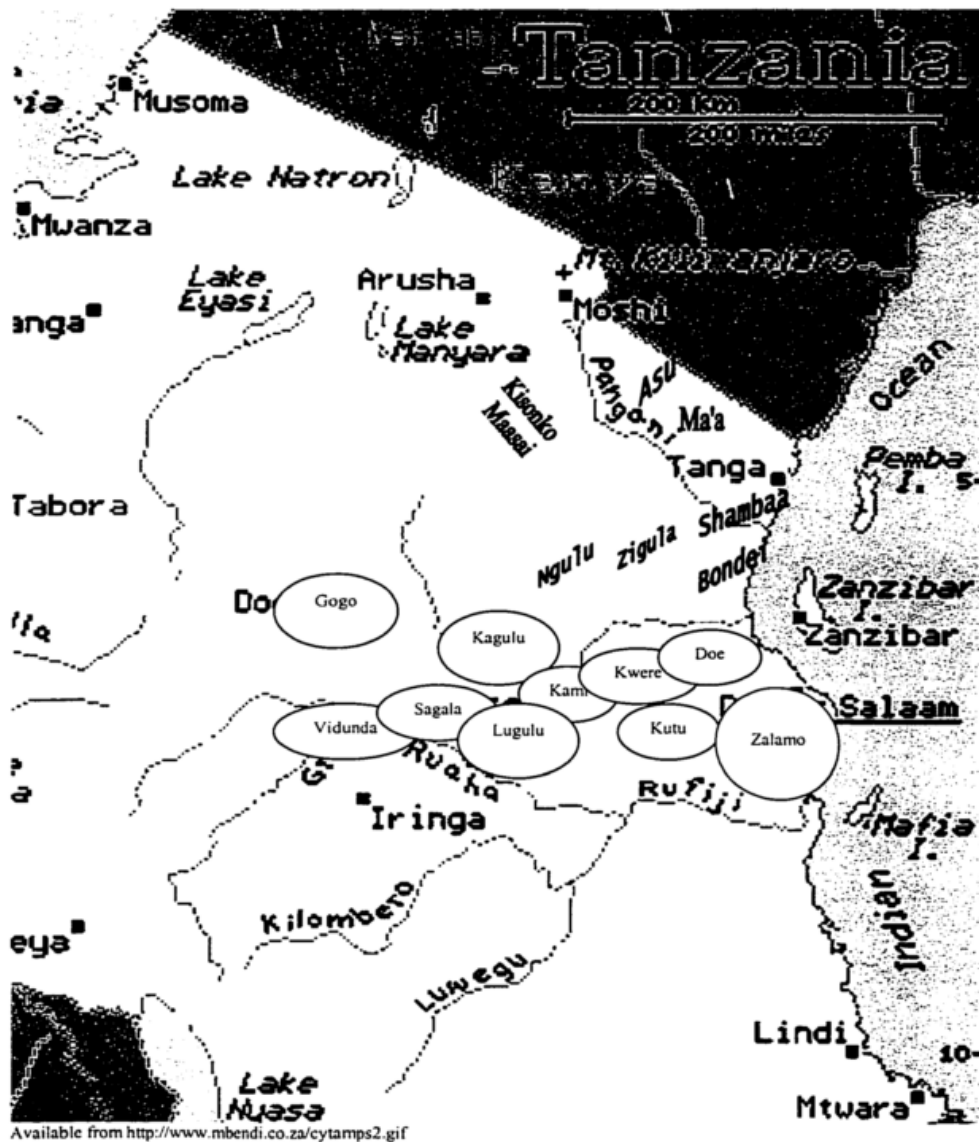
Figure 1 - Map situating the Sagala and the other Ruvu peoples[2], c. 1700-1900 A.D. (Gonzales 2002: 53)

For the purpose of situating Sagala in its language family, multiple classifications need to be discussed. Firstly, there is Guthrie's classification of the Bantu languages, a referential system based on phonological, grammatical and lexical criteria (Polomé 1980: 16). Guthrie included languages in his classification following two principal (1-2) and two subsidiary (3-4) criteria: (1) a system of grammatical genders (or noun classes), usually at least five; (2) a vocabulary, part of which can be related by fixed rules to a set of hypothetical common roots;

---

[2] The size and shape of the symbols used are not indicative of population size (cfr. Gonzales 2002: 42).

(3) a set of invariable cores, or radicals, from which a majority of the words are formed by an agglutinative process; (4) balanced vowel system in the radicals, consisting of one open vowel 'a' with an equal number of back and front vowels (Guthrie 1967: 11-12). Using these criteria Guthrie (1967) divided the Bantu languages into 16 geographical zones, each of which were given a letter. Within these zones groups of languages were indicated by a figure, and the number of an individual language within a group by a second figure (Guthrie 1967: 29). Guthrie labelled Sagala with the code G39 (Guthrie 1967: 48), which means that Sagala was the ninth language of the third group of zone G. Maho (2009) has kept this code in his *New Updated Guthrie List*, in which he called G30 the 'Zigula-Zaramo Group', in reference to Zigula [G31] and Zaramo [G33]. For the remainder of the paper, languages will be labelled according to Maho's updated version of Guthrie's divisions (cf. Petzell & Hammarström 2013: 129).

The Nurse and Philippson (1980) classification of the Bantu languages of East Africa is lexicostatistical. They group languages based on shared cognates in a wordlist of 400 concepts. Petzell and Hammarström explain their method as follows:

> "Once lexicostatistical percentages for each pair of languages have been calculated, the languages are broadly classified into groups within which the average percentage of similarity is higher within the group than in comparison with the most similar language outside the group ("strong groups") or almost so ("weak groups")." (Petzell & Hammarström 2013: 132)

Nurse and Philippson (1980) classify Sagala historically into a West Ruvu group together with Gogo [G11] and Kagulu [G12], as is shown in Figure 2[3].

---

[3] In the figure, 'Sagala' is written as 'Sagulu'. I assume this is simply a typing error, as in previous pages of their book, Nurse and Philippson do speak of 'Sagala' as being part of this group.
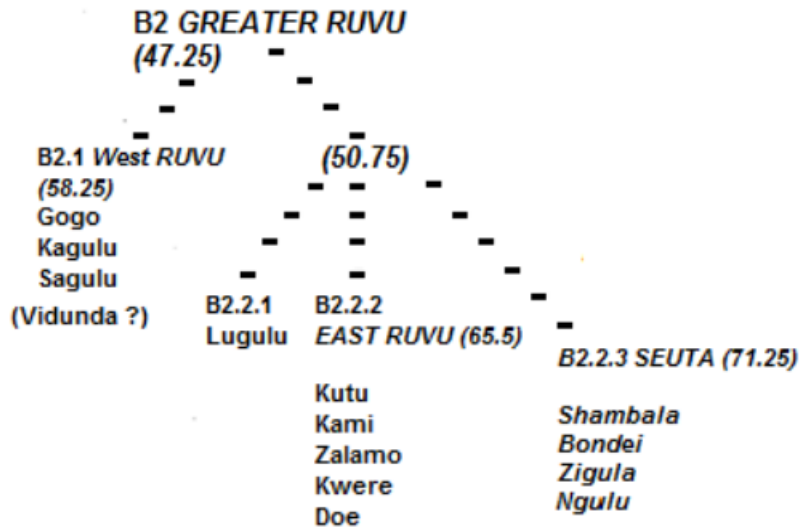
Figure 2 - Nurse and Philippson's (1980: 50) classification of the Greater Ruvu languages (also presented in Petzell & Hammarström 2013: 132)

The last classification I discuss is that of Gonzales (2002), shown in Figure 3. She uses a methodology that is similar to the one of Nurse and Philippson, except that she uses only 100 words and includes potential borrowings (Petzell & Hammarström 2013: 132). Following Nurse and Philippson, she claims that Sagala belongs to a West Ruvu group. The difference, however, is that Kagulu is not included in this group, whereas Vidunda [G38] is (Gonzales 2002: 38). Gonzales' classification also shows that Sagala is most closely related to Vidunda.
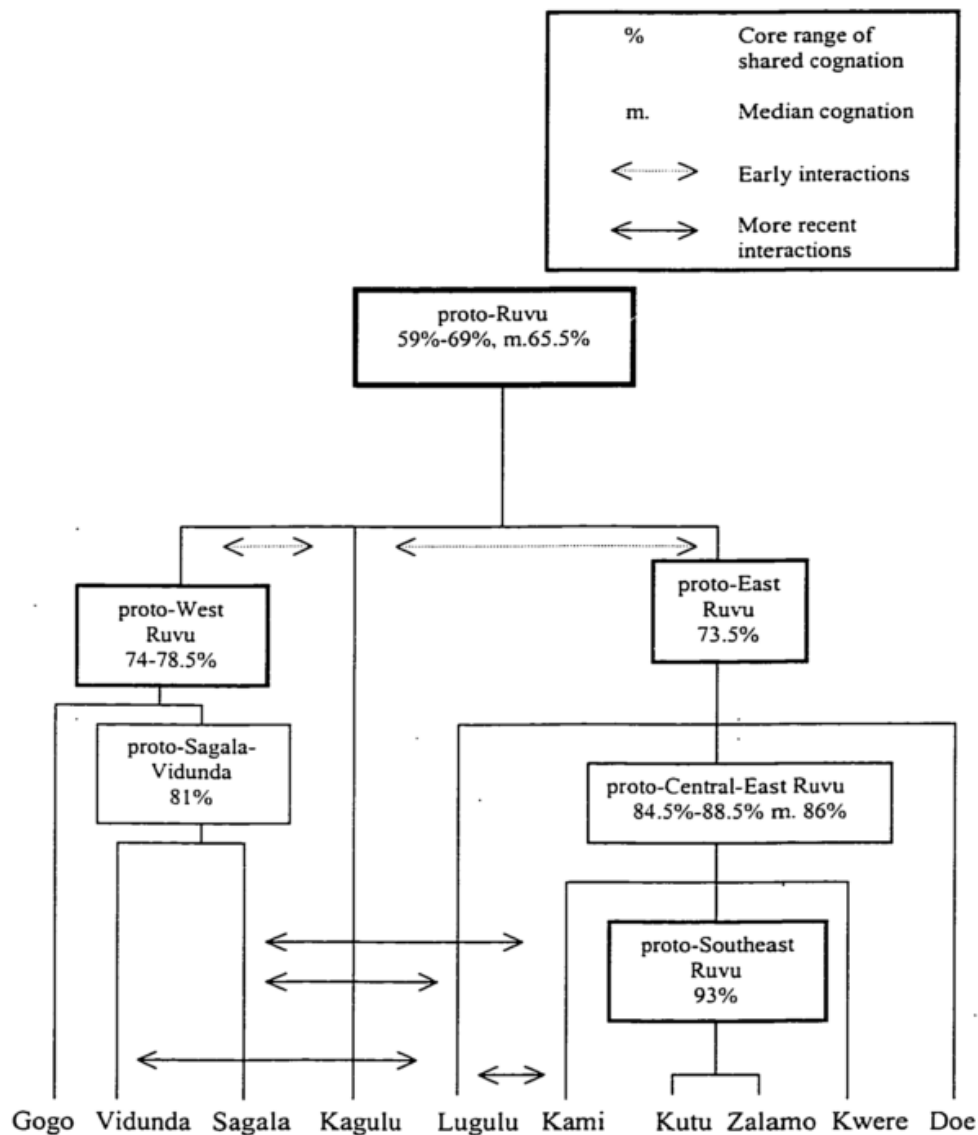
Figure 3 - Gonzales' (2002: 34) classification of the Ruvu languages (also presented in Petzell & Hammarström 2013: 134)

1.2.   State of the art and motivation for research

Beidelman wrote in 1967 that

> "The matrilineal peoples of eastern Tanzania are among the most poorly described societies of East Africa, even though they are among those with the earliest and longest contact with alien, literate societies." (Beidelman, as cited in Gonzales 2002: 11)

Similarly, Petzell and Hammarström (2013: 129) write that the language varieties they consider as the Greater Ruvu Bantu languages, to which Sagala is closely related, are poorly described. Both claims also hold for Sagala.

Especially when it comes to linguistic documentation and description, it seems very little material is available. Prof. Karsten Legère (formerly at Gothenburg University) made a number of recordings of Sagala discourse which he kindly shared with us for future research. Some of these have been transcribed, but they were not exploited yet for linguistic research.

Dr. Malin Petzell (Gothenburg University) did research on Kagulu [G12] and on several members of the G30 group, except Vidunda and Sagala. She has published several works, of which *Grammatical and Lexical Comparison of the Greater Ruvu Bantu Languages* (2013) (which has Harald Hammarström as co-author) and *The Kagulu Language: Grammar, Texts and Vocabulary* (2008) have proved to be of great help to me. Apart from those languages, Vidunda has also been described to some degree. Legère has worked on it, focusing on endangerment and its use of plant names, among other things, see *Vidunda People and their Plant Names* (2004), *Plant Names in the Tanzanian Bantu Language Vidunda: Structure and (Some) Etymology* (2009) and *Vidunda (G38) as an Endangered Language?* (2007).

It seems that Sagala is the only G30 language which has remained entirely non-described. It is important to change this. Language documentation has been recognised as a way to safeguard indigenous languages, and endangered languages in particular (Batibo 2009: 193). Batibo states that

> "(…) a systematic description and codification of the indigenous languages would empower these languages for public use, preserve them for future generation as well as give them more utilitarian value." (Batibo 2009: 193)

As it is, I am not certain whether Sagala really is an endangered language. To assess whether or not a language is endangered, or the degree to which it is, the UNESCO criteria prove most useful (UNESCO 2003). One of these criteria is the absolute number of speakers. UNESCO, however, claims it "impossible to provide a valid interpretation of absolute numbers" (UNESCO 2003: 8), but stresses that "a small speech community is always at risk" (UNESCO 2003: 8). As I have mentioned earlier, in 2008, Sagala was said to have around 106,000 speakers. However, although this might be considered a reasonable amount of speakers compared to the smallest languages, which do not even reach 1000 speakers, Muzale and Rugemalira stress that "The level of language endangerment for even the most populous speech communities is considerable given the ever-rising fortunes of Kiswahili" (Muzale & Rugemalira 2008: 84).

Swahili, Tanzania's *lingua franca* or language of wider distribution, is a real threat to the other languages of Tanzania. It has had a major influence, causing gradual marginalisation of smaller languages (Legère 2007: 43). This is partly due to the use of Swahili as the medium of instruction in primary schools, which has a strong impact on the younger

generation's competency and proficiency of the smaller languages. Moreover, smaller languages are excluded from most formal domains, where Swahili plays a prominent role as the national and co-official language. Lastly, after independence in 1961, the existence of languages other than Swahili was ignored and/or considered a breeding ground for tribalism. This may also have contributed to the shift to Swahili (Legère 2006: 99-100, 107). Legère has written more on language endangerment in Tanzania, see *Language endangerment in Tanzania: identifying and maintaining endangered languages* (2006) and *Vidunda (G38) as an Endangered Language?* (2007).

The dominance of Swahili may lead to a deterioration of intergenerational language transmission, which is another one of UNESCO's criteria. However, my research has been very small-scale and I have not made it a focus to decide whether Sagala is an endangered language. Consequentially, I do not have the necessary information to do so. I will therefore keep Muzale and Rugemalira's statement in mind that all language communities in Tanzania are to some degree endangered, due to the rise of Swahili. The threat of marginalisation because of influence of Swahili in itself legitimises the documentation of Sagala. Moreover, every language deserves to be documented, in an attempt to hold onto the cultural knowledge embedded in it, and to help future linguistic research. The purpose of this paper is to be of some help in this particular endeavour.

1.3.   Methodology

When it comes to languages which have not yet been the subject of linguistic research, language documentation is in order before language description becomes possible. These two activities, language documentation and language description, differ in both their methods and their results. The former consists of the collection, transcription and translation of primary data, while the latter refers to a descriptive analysis of those data (Himmelmann 1998: 161). For my research on Sagala, I have undertaken both activities.

The first phase, documentation, was executed during a five-month stay in Morogoro, Tanzania in 2016-2017. The purpose of this stay was twofold: attending courses at Mzumbe University for the duration of one semester as well as doing fieldwork for this paper. I have conducted my fieldwork in the centre of Kilosa, a town 115 km west of Morogoro. There, I got to meet a few Sagala-speakers, three of which ultimately became my consultants. They are called Juma, Amina and Ndoweka. An identification sheet is provided in Appendix B, but I will provide some information about each of them here.

Juma was born in 1942 in Mkadage, a village in Kilosa district where he also went to school. When he finished school, he went to Dar es Salaam where he served in the National Service

and in the Tanzania People's Defence forces. During these years he has visited several places. In 1965 he went to Tel Aviv, Israel for six months to study agriculture and settlement. He married in 1967 and has seven children. Nowadays, he is a farmer and lives close to the centre of Kilosa. His mother tongue is Sagala, although he mostly used Swahili in everyday communication. Because of his years as a soldier, he has a basic knowledge of Sukuma.

Amina was born in 1957 in Munisagala, another village in Kilosa district. She finished school in 1971 after which she went to Dodoma where she married. Throughout her life, she has lived in Dodoma, Mwanza, Iringa, Morogoro and Moshi, as well as Kilosa. Her main reason for going to these places was trade. Amina has five children. Her first language is Sagala, but she has also spoken Swahili since she was young. She used to speak Gogo when she lived in Dodoma but only very little of this knowledge remains.

Lastly, Ndoweka is 86 years. He was born in the village of Tame, in Kilosa district. He has lived in multiple villages in Kilosa district. For his studies he went to Kenya, where he has returned multiple times for work obligations. He has also worked in Kilosa, mostly in accountancy. He has been to Dodoma, Dar es Salaam, the Kilimanjaro region, Morogoro and Mahenge. He is retired and lives in Kilosa. Ndoweka's first language is Kisagala, but he also speaks Kaguru, Vidunda, Luguru and Kwere. He says these languages are very alike and share a considerable amount of words. Other than that, Ndoweka speaks Swahili and some English. His father spoke Sagala and Luguru, while his mother spoke Sagala and Kaguru. At home, they spoke Zungwa, which is the variety of Sagala both of his parents spoke. Swahili was used in everyday communication outside the house.

As a sort of preparation for my fieldwork, I have participated in the *11th Summer School in Languages and Linguistics* in Leiden. I followed the workshops on Descriptive Linguistics, Tone Analysis, and Field Methods. The latter, taught by Dr. Christian Rapold, was especially useful for my purposes. The knowledge acquired in this workshop has helped me a lot while I was in Tanzania, as has Claire Bowern's *Linguistic Fieldwork: A Practical Guide* (2008), which was recommended during the workshop.

The gathering of data during the documentation stage may be achieved in multiple ways, such as participant observation and elicitation (Himmelmann 1998: 162). I have employed elicitation, whereby I had a prepared wordlist at my disposal and asked the speakers to translate these words to the best of their abilities. I used the 500-wordlist Petzell and Hammarström used in their study of the Greater Ruvu languages (2013). In that way, the data and results achieved through this research can facilitate future research on this group of languages.

Given that elicitation is a form of interviewing, it is not a natural setting for using a language. This was further reinforced by the fact that two consultants, Juma and Amina, with whom I had joined sessions, preferred to obtain a copy of the wordlist beforehand, in order for them to be able to prepare it at home. As this made them much more comfortable in their abilities, I opted to continue like this after a few sessions.

The sessions were almost all recorded, except when the speakers said they were not comfortable with it. This only happened twice. The consultants were asked to translate words into Sagala. Even though with Juma and Amina the medium of communication was mostly English, the words were asked in Swahili. With Ndoweka the medium of communication was Swahili, but a middleman was present most of the time to assist when Ndoweka and I did not understand each other.

When dealing with nouns, it proved useful to ask the plural in addition to the singular. This slipped my mind a couple of times, and it has given me some trouble during the analysis of the data. Verbs were sometimes offered in the infinitive, sometimes in the imperative mood, depending on what felt most natural to the speaker. Adjectives were supposed to be paired with multiple nouns, but due to a lack of time, most were only paired with one. Upon asking, Amina, who seemed to be the most proficient speaker, used a significant amount of the vocabulary in sentences. The consultants were also told they were free to add any and all supplementary information, which led to a small amount of additional data, such as proverbs and small bits of songs. The words elicited in isolation have been transcribed and added to the wordlist, which can be found in Appendix A. The vocabulary is available in English, Swahili and Sagala.

After the data have been collected, transcribed and translated, the necessary material is available to start with the second phase, description. The results of that analysis are presented in the following sections. There are two sections, the first treats the synchronic phonology of Sagala and the second elaborates on the nominal morphology of the language.

1.4.   Orthography

Orthography can be defined as "the *conjunction* of a set of graphemes, such as an alphabet, and a set of accompanying rules regulating their use" (Seifart 2006: 277, italics in original). Both the symbols and their usage are codified (Seifart 2006: 277). Orthography is a matter of concern in any attempt at language documentation and/or description. Without fail, the researcher will have to make decisions on this front.

There is not a single orthography for all African languages, which should not come as a surprise, given that the variation in sounds among the African languages is immense. Even if one narrows down to Tanzania, there still is not a common orthography for all languages spoken there, mainly for the same reason. Many languages in Tanzania, including Sagala, merely exist in oral form without written records. As a consequence, the necessity to develop an orthographic system is an additional issue for researchers working on such languages. In doing so, multiple things need to be considered.

First and foremost, the researcher needs to be aware that the orthography they develop should be available to be used by the speakers of the language, who would benefit the most from the ability to write down their language. They do not need phonetic assistance for correct pronunciation, something the researcher needs to take into consideration, even though they might want to use a more phonetic representation for scientific reasons (Meinhof & Jones 1928: 233). Secondly, the speakers probably know or have some notion of the orthography of a vehicular language that has been put into writing, or of a European language they are familiar with (Meinhof & Jones 1928: 235-236). Therefore, the newly developed orthography should not differ too substantially from the one they already know, or at the least have notion of, to avoid making things unnecessarily complicated.

In my decision on how to put Sagala into writing, a process that will always be somewhat arbitrary, previous works on related languages have been of great use. Legère (2004: 118) notes that Vidunda has not been reduced to writing. He therefore uses a slightly modified version of the Swahili orthography. Furthermore, Petzell (2003: 1) says that the Kagulu orthography is a somewhat modified version of the Swahili writing convention. Since Sagala has specific sounds that are not present in Swahili but are shared with Vidunda and Kagulu, I feel justified in opting to get my inspiration from the orthographies Legère and Petzell used for them. Throughout the paper, I will elaborate further on specific orthographic matters in the paragraphs in question.

The orthography which I use is neither a deep orthography nor a shallow one. A shallow orthography would represent a highly phonetic realisation of linguistic forms (Seifart 2006: 279). However, every town and every village has its own regiolect of Sagala. Even individuals speak differently, something I noticed in the speech of my consultants. If I were to use a shallow – phonetic – orthography, it would only show the pronunciation of a village or of an individual (Meinhof & Jones 1928: 230). A deep orthography, on the other hand, represents underlying forms and does not specify morphophonological changes (Seifart 2006: 279). Seeing as I do specify some morphophonological changes, but refrain from writing entirely phonetically, I would say my orthography of choice is somewhere in between.

A list of the graphemes is provided in Table 1. The International Phonestic Alphabet (IPA) is used for the phonemes. There is an example for each phoneme both in word-initial (if available) and word-medial position (cf. Petzell 2008: 36-38).

Table 1 - Graphemes and phonemes in Sagala

Grapheme	Phoneme	Example	Translation	Comment
a	a	*nzala*	'hunger'	
b	b	*bangu, ibakwa*	'scar, door'	
ch	c	*chigoda, machifu*	'chair, ashes'	Only used by Ndoweka.
d	d	*diziko, ngodi*	'fireplace, firewood'	
e	ɛ	*mfele*	'liver, woman'	This phoneme can also be realised as [e].
f	f	*finzi, lufu*	'darkness, intestine'	
g	g	*guha, fagilo*	'bone, broom'	
h	h	*hegulo, ibahu*	'morning, year'	
i	i	*ivuke, dizigo*	'heat, load'	
j	ɟ	*jamvi, meji*	'mat, water'	This phoneme can also be realised as [dj]. It was only used by Ndoweka.
k	k	*kutu, nzuki*	'cheek, honey bee'	
l	l	*lukolo, palati*	'clan, buffalo'	
m	m	*mkala, kipehema*	'charcoal, chest'	
mb	^{m}b	*mbeyu, dihamba*	'seed, leaf'	
mh	m^{h}	*mhene*	'goat'	
n	n	*nemisi, mwana*	'daytime, child'	
nd	^{n}d	*ndoto, mgunda*	'basket, field'	
ng'	ŋ	*ng'ombe, kung'ala*	'cattle, to shine'	
ng	^{ŋ}g	*ngedele, gongo*	'monkey, mountain'	
ng'h	ŋ^{h}	*ng'handa, kunung'ha*	'guinea fowl, to stink'	
nh	n^{h}	*nhembo, munhu*	'elephant, person'	
ny	ɲ	*nyakongo, kinyala*	'soldier ant, shame'	
nz	^{n}z	*nzila, mgenzi*	'path, guest'	
o	ɔ	*kombe*	'fingernail'	This phoneme can also be realised as [o].

p	p	*p*aga, ng'hwa*p*a	'rib, armpit'	
s	s	*s*akame, m*s*ile	'blood, arrow'	
t	t	*t*ombo, gu*t*i	'breast, ear'	
u	u	*u*mbagi, mt*u*mbwi	'spear, canoe'	
v	v	*v*agi, m*v*ula	'quarrel, rain'	
w	w	*w*engu, dita*w*i	'lung, branch'	
y	j	*y*ega, ma*y*a	'shoulder, anger'	
z	z	*z*iso, di*z*oka	'eye, snake'	

2. Synchronic phonology

2.1.   Introduction

Before presenting a phonological analysis of Sagala, let us first define what exactly that term, 'phonology', entails. According to Lass it is

> "that subdiscipline within linguistics concerned with 'the sounds of language'. More narrowly, phonology proper is concerned with the function, behaviour, and organization of sounds as linguistic items; as opposed to phonetics, which is a rather more 'neutral' study of the sounds themselves as phenomena in the physical world, (…)." (Lass 1984: 1)

Central to the analysis of a language's phonology are its basic speech sounds, viz. phonemes. Those are minimal units that serve to distinguish words from each other. Minimal pairs are the most effective way to prove that two sounds are distinct phonemes (Hayes 2009: 32-33). These are sets of words that differ by a single phoneme (Barlow & Gierut 2002: 58). One phoneme, in its turn, may have multiple phonetical realisations. These are called allophones, and these may be either free or tied to a certain environment. If the latter is the case, and one sound never occurs in the environment in which the other occurs, allophones are said to be in complementary distribution (Hayes 2009: 21).

When it comes to the phonological structure of a language, especially a Bantu language, there are two levels to be taken into account, i.e. the segmental and the supra-segmental. 'Segmental' refers to the phonemes, whereas 'supra-segmental' refers to the prosodic features of a language, such as tone, stress and intonation (Soto-Faraco & Sebastián-Gallés 2001: 412). I will start this section about phonology with the synchronic segmental phonology of Sagala by discussing vowels, consonants and glides. Then I will treat syllable structure in Sagala, continue with prosody, and end with certain morphophonological processes taking place. It is important to note in advance that I have not recognised Sagala as a tone language, leading me to not mark tone in this paper. I will elaborate on this issue further down, in paragraph 2.6.

2.2.   Vowels

Vowels are different from consonants in that they do not have a place of articulation. This means that there is no point of major constriction in the vocal tract, which is the term used to denote the portions of the human anatomy through which air flows in the course of speech production (Hayes 2009: 1, 11).

Vowels are described according to the following three criteria: rounding, height and backness. Through rounding of the lips, the passage at the exit of the vocal tract is narrowed. Secondly, the height of the tongue, accompanied by either opening or raising the jaw, makes the passage through the mouth wider or narrower. Lastly, the body of the tongue may be placed either towards the front or towards the back of the mouth (Hayes 2009: 11-12).

Table 2 - Vowel phonemes in Sagala

Backness / Height	Front	Central	Back
High	i		u
(Open-)mid	ɛ		ɔ
Low		a	

The mid vowel phonemes /ɛ/ and /ɔ/ will be orthographically presented as <e> and <o>.

Sagala is a 5V language, meaning it has five vowel phonemes presented in Table 2. There are the high vowels /i/ and /u/, the open-mid vowels /ɛ/ and /ɔ/ and the low vowel /a/. The central and the front vowels are unrounded and the back vowels are rounded. Phonetically speaking, the mid vowel phonemes can be realised as both the close-mid/half-close vowels [e] and [o] and the open-mid/half-open vowels [ɛ] and [ɔ]. I observed free allophonic variation, depending on the speaker, between both phonetic realisations.

I will shortly explain why I consider [ɛ] and [ɔ] as the basic realisations of Sagala's mid vowel phonemes. Firstly, while Proto-Bantu had the 7V system presented in (1a) (Bostoen 2008: 307), a large part of Bantu languages have merged the first and second degree vowels to achieve the 5V system as shown in (1b). It is rare for current 5V Bantu languages to have [e] and [o] instead of [ɛ] and [ɔ] (Hyman 2003: 45), which is one of the reasons to lean in favour of the latter.

(1)   (a) 7V   (b) 5V

          *i   *u   i   u
          *ɪ   *ʊ   ɛ   ɔ
          *ɛ   *ɔ   a
              *a

Secondly, I often had trouble distinguishing between [e] or [ɛ], and [o] or [ɔ] in certain words. The quality of the vowels became clearer, however, when I asked the consultants to repeat once or twice. When they spoke more slowly and they articulated more clearly, the mid-open vowels [ɛ] and [ɔ] were easier to perceive.

The high front vowel /u/ was often slightly nasalised when following a bilabial nasal /m/ or the aspirated nasals /nh/ and /ng'h/ (cf. infra).

2.2.1. Minimal pairs

For each vowel multiple minimal pairs are presented in the following list. The contrasting sounds are underlined. The translation of the words is given in English. For a translation in Swahili, check the wordlist in Appendix A. This will continue to be the case for the rest of the paper.

(2)	/i/ - /ɛ/	kul_i_la 'to cry'	kul_e_la 'to look after a child'
	/ɔ/ - /a/	kul_o_nga 'to speak'	kul_a_nga 'to show'
	/ɛ/ - /ɔ/	kus_e_ka 'to laugh'	kus_o_ka 'to be tired'
	/ɔ/ - /u/	kul_o_ta 'to dream'	kul_u_ta 'to go'
	/i/ - /ɔ/	ml_i_mo 'work'	ml_o_mo 'mouth'
	/u/ - /ɛ/	kut_u_nda 'to urinate'	kut_e_nda 'to do'
		kus_u_ka 'to plait'	kus_e_ka 'to laugh'
	/a/ - /u/	md_a_la 'old person'	md_u_la 'hyena'
		kug_a_la 'to bring'	kug_u_la 'to buy'
		kus_a_ka 'to hunt'	kus_u_ka 'to plait'

2.2.2. Vowel length

Hyman (2003: 48) mentions five sources of vowel length in Bantu, three of which occur in Sagala. The first of these is compensatory vowel lengthening as a consequence of gliding. As I will discuss below (cf. § 2.7.1.), gliding means that the front vowels /i, e/ are realised as [y] when followed by a non-identical vowel, and the back vowels /u, o/ as [w]. This can happen when two vowels belonging to different morphemes come into contact. Gliding accompanied by vowel lengthening is exemplified in (3). The lengthened vowels are in a NGV or GV positon. The second source Hyman mentions that counts for Sagala is vowel

lengthening before a moraic nasal + consonant. Examples are found in (4). Here, the lengthened vowels are in a VNC position. Lastly, Hyman mentions penultimate vowel lengthening. He says that it "occurs in most eastern and southern Bantu languages which have lost the lexical vowel length contrast, (…)" (Hyman 2003: 48). Seeing as some of the previous examples also apply to this criterium, I give a few examples (5) where the vowel is neither in a NGV/GV nor a VNC position.

In each of these contexts, the phonetic lengthening of the vowel is predictable. Phonemic vowel length does not occur in Sagala.

(3)	*[tate] yetu* 'our [father]'	GVCV
	mwana 'child'	NGVNV
(4)	*kukonga* 'to begin'	CVCVNCV
	kutenda 'to do'	CVCVNCV
	nhungika 'to hang up'	CVNCVCV
(5)	*kusogota* 'to twist (rope)'	CVCVCVCV
	kubena 'to break (tr.)'	CVCVCV

2.3.   Consonants

As mentioned above, there needs to be some kind of constriction in the vocal tract in order to form consonants. As vowels, consonants are also classified along three criteria, i.e. voicing, place of articulation and manner of articulation. The first criterium denotes whether or not the vocal cords vibrate (Hayes 2009: 6). Next, there are various places and manners of articulation. I will shortly explain those which apply to Sagala.

Place of articulation refers to where in the vocal tract the constriction occurs. Bilabial sounds are formed by touching the lips together. Labiodental sounds are formed by touching the lower lip to the upper teeth. Alveolar sounds are formed by touching the tip of the tongue to a location just forward of the alveolar ridge. Palatal sounds are formed by touching the tongue blade and the forward part of the tongue body to the hard palate. Velar sounds are formed by touching the body of the tongue to the hard or soft palate. Glottal sounds are formed by moving the vocal cords close to one another. Lastly, labial-velar sounds are formed by both bringing the lips together and touching the body of the tongue to the soft palate (Hayes 2009: 7-8).
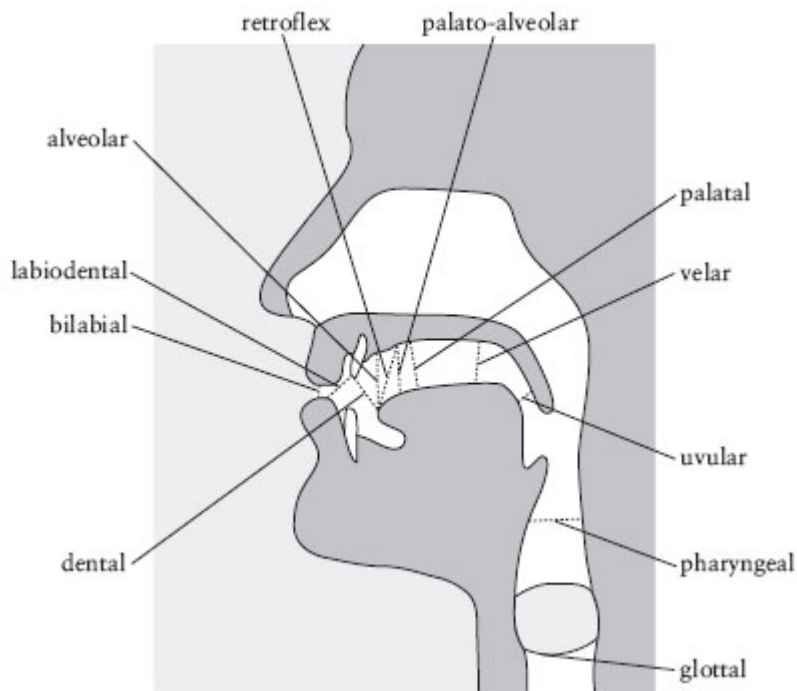
Figure 4 - Places of articulation

Then, there are the different manners of articulation. These may all occur in the different places of articulation that were just explained. In a stop the airflow through the mouth is momentarily closed off. For a nasal, the velum is lowered, causing the air to escape through the nose. In a fricative, a tight constriction is made in the vocal tract, so that the air passing through it makes a hissing noise. In approximants, the constriction is fairly wide, so that the air passes without creating turbulence. In particular for lateral approximants, the air passes around the sides of the tongue (Hayes 2009: 6-7). Both stops and fricatives may be prenasalised. This phenomenon is explained further in the paper (cf. § 2.7.3).

Table 3 - Consonant phonemes

Place of artic. / Manner of artic.	Bilabial		Labio-dental		Alveolar		Palatal		Velar		Glottal	Labial-velar
Stop	p	b			t	d	ch	j	k	g		
Nasal		m				n		ny		ng'		
Fricative	f	v	s	z							h	
Approximant								y				w

Lateral approximant			l			
Prenasalised stop	mb		nd		ng	
Prenasalised fricative			nz			

In the table I have used the orthography that is common for Bantu languages. In (6) I present the phonetic symbols of those sounds for which I use a different orthography.

(6)   Orthography   IPA

                     <ch>   [c]

                     <j>   [ɟ]

                     <ny>   [ɲ]

                     <ng'>   [ŋ]

                     <y>   [j]

                     <ng>   [ŋg]

It should also be noted that /ch/ and /j/ only occurred in sessions with Ndoweka. The other consultants used multiple alternatives for this depending on the context or the word they were used in. For example, Ndoweka used /chi/ as nominal class prefix for class 7 (cf. infra), while Juma and Amina used /ki/. For 'ashes,' Ndoweka gave *machifu*, while the offers gave *madivu*. For 'to boil (intr.),' Ndoweka gave *kuchemsa*, while the others gave *kuhemsa*. When it comes to the second sound, /j/, Juma and Amina either used /z/ instead, e.g. *kwiza* while Ndoweka gave *kwija*, or they gave translations that did not resemble Ndoweka's. /j/ was also often phonetically realised as [dj].

Voiceless stops may be slightly aspirated, depending on the speaker. This was very common in the speech of Juma and Ndoweka. Aspiration is not contrastive in this context.

2.3.1. Minimal pairs

For all the above-mentioned consonants, at least one minimal pair is given in (7). As was the case with the vowels, the contrasting sounds are underlined.

(7) /v/ - /w/   *kulava [nyasa]* 'to sneeze'   *kulawa* 'to come from'

/l/ - /g/   *kulila* 'to cry'   *kuliga* 'to insult'

/b/ - /l/   *kubawa* 'to gather'   *kulawa* 'to come from'

/b/ - /g/   *kubela* 'to lie'   *kugela* 'to put into'

/l/ - /t/   *kulola* 'to see'   *kulota* 'to dream'

/n/ - /l/   *kugona* 'to sleep'   *kugola* 'to harvest'

/t/ - /n/   *kubeta* 'to wait'   *kubena* 'to break (tr.)'

/s/ - /k/   *kusola* 'to take'   *kukola* 'to seize'

/nd/ - /z/   *kuhinda* 'to shut'   *kuhiza* 'to steal'

/z/ - /l/   *kuguza* 'to sell'   *kugula* 'to buy'

/nd/ - /l/   *kuhonda* 'to smash'   *kuhola* 'to cool down'

/h/ - /l/   *kuliha* 'to pay'   *kulila* 'to cry'

/l/ - /k/   *kulema* 'to refuse'   *kukema* 'to call'

/l/ - /ng/   *kulola* 'to see'   *kulonga* 'to speak'

/m/ - /l/   *kuhoma* 'to burn up'   *kuhola* 'to cool down'

/ny/ - /l/   *kumanya* 'to know'   *kumala* 'to finish (tr.)'

/g/ - /t/   *kugenda* 'to walk'   *kutenda* 'to do'

/m/ - /k/   *kuwima* 'to stand'   *kuwika* 'to lay down'

/t/ - /s/   *kutaga* 'to throw'   *kusaga* 'to grind'

/s/ - /l/   *kusola* 'to take'   *kulola* 'to see'

/t/ - /m/   *kuluta* 'to go'   *kuluma* 'to bite'

/g/ - /m/   *kugala* 'to bring'   *kumala* 'to finish (tr.)'

/z/ - /mb/   *kuhiza* 'to steal'   *kuhimba* 'to dig'

/nz/ - /nd/   *kuhanza* 'to mix'   *kuhanda* 'to plant'

/nd/ - /g/   *kufunda* 'to learn'   *kufuga* 'to rear'

/nd/ - /mb/   *kuhinda* 'to shut'   *kuhimba* 'to dig'

/m/ - /ng/	*kuwima* 'to stand'	*kuwinga* 'to chase away'
/k/ - /ng/	*kuwika* 'to lay down'	*kuwinga* 'to chase away'
/k/ - /n/	*kuvina* 'to play'	*kuvika* 'to dress (tr.)'
/nz/ - /mb/	*kuhanza* 'to mix'	*kuhamba* 'to decorate'
/s/ - /d/	*kuseka* 'to laugh'	*kudeka* 'to vomit'
/v/ - /nz/	*kulava [nyasa]* 'to sneeze'	*kulanza* 'to ask for'
/w/ - /nz/	*kulawa* 'to come from'	*kulanza* 'to ask for'
/ng/ - /nd/	*ng'hanga* 'guinea fowl'	*(i)ng'handa* 'house'
/b/ - /y/	*bangu* 'scar'	*[tate] yangu* 'my [father]'
/f/ - /t/	*kufunda* 'to learn'	*kutunda* 'to urinate'
/k/ - /ng'/	*kombe* 'fingernail'	*ng'ombe* 'cattle'
/m/ - /ng'/	*kumala* 'to finish (tr.)'	*kung'ala* 'to shine'
/g/ - /h/	*kugona* 'to sleep'	*kuhona* 'to sew'
/nd/ - /n/	*kuhonda* 'to smash'	*kuhona* to sew'
/p/ -/s/	*paga* 'rib'	*saga* 'grind'
/v/ - /d/	*mvula* 'rain'	*mdula* 'hyena'
/w/ - /y/	*wake* 'women'	*[tate] yake* 'your [father]'
/ch/ - /f/	*chidole* 'finger'	*fidole* 'fingers'

2.3.2. Allophone [r]

As is the case in many, if not all, East-African languages (Schadeberg 2009: 89), [l] and [r] are free allophones in Sagala. I have identified [l] as the basic phonetic realisation of this liquid phoneme. Free allophonic variation is possible, whereby speakers use [r] instead of [l]. The consultants did not make a difference between a word being pronounced with [l] and the same word being pronounced with [r]. They interchanged the two frequently, e.g. [ladu] 'lightning' or [radu], and [lwanda] 'river' or [rwanda]. In the wordlist in Appendix A, some words are represented with [r], when this was the only version of the word heard in the elicitation sessions.

2.4. *Glides*

There are multiple opinions on how to define glides, saying that in se they are vowels, or consonants, or not completely either one. They are produced with a relatively unimpeded flow of air through the mouth, likening them more to vowels, but then, they might also function phonologically as consonants, placed on the edge of a syllable (Zhang 2006: 105). Sagala has two glides, /y/ and /w/. They can be found under 'Approximants' in the table of consonants (cf. supra).

2.5. *Syllable structure*

In every language, there are rules determining the combination of phonemes, which leads to every language having its canonical syllable structure. For many African languages, this is CV. This would mean that there cannot be any consecutive consonants within the same syllable, although there are, of course, exceptions (Meeuwis & Gunnink 2014: 148).

In Sagala most syllables are open, but syllables consisting of a single nasal are also allowed. This leads to the possible syllable structures being: CV, V, CGV, N, NV, NCV, NGV, NCGV, GV. Examples for each are presented in (8).

(8)	*ziso* 'eye'	CV-CV
	mlomo 'mouth'	N-CV-CV
	tagwa 'name'	CV-CGV
	ngodi 'firewood'	NCV-CV
	mwana 'child'	NGV-NV
	nswa 'termite'	NCGV
	yega 'shoulder'	GV-CV
	umbagi 'spear'	V-NCV-CV

There is no maximum number of syllables. As is clear, there are instances where multiple consonants are combined in the same syllable. I will now take a closer look at each of these.

2.5.1. N + C combinations

NC-combinations are stops or fricatives preceded by a nasal. A distinction between two kinds is to be made based on place of articulation. If the NC-combination is homorganic, it means that the nasal is articulated at the same place as the stop or the fricative, e.g. a bilabial NC-combination /mb/. Homorganic NC-combinations can be the result of a prenasalisation of the consonant (cf. infra), leading to one phoneme, or they can be a succession of two separate phonemes. Heterorganic NC-combinations, the second kind, are also a succession of two phonemes, but these differ in place of articulation (Meeuwis & Gunnink 2014: 144-145). Since they cannot belong to the same syllable, heterorganic NC-combinations are of no importance when discussing the canonical syllable structure of Sagala. Therefore, they will not be treated further.

The possible NC-combinations in Sagala are presented in (9). They all consist of a prenasalised consonant, therefore the NC-combinations are homorganic. The NC-combinations in question are underlined.

(9)   N + /d/ → /nd/ [nd]   *lwanda* 'river'

        N + /z/ → /nz/ [nz]   *ganza* 'palm (of hand)'

        N + /s/ → /ns/ [ns]   *nzagansa* 'lip'

        N + /b/ → /mb/ [mb]   *tombo* 'breast'

        N + /g/ → /ng/ [ŋg]   *kuzenga* 'to build'

2.5.2. C + G combinations

A consonant may be followed by a semivowel (10). This is possible since the latter has the physical characteristics of a vowel, even though it acts like a consonant, as mentioned earlier.

(10)   /p/ + /w/   *ipwa* 'to ripen (intr.)'

        /b/ + /w/   *kibwengu* 'ankle'

        /t/ + /w/   *ditwi* 'head'

        /l/ + /w/   *lwangi* 'voice'

        /k/ + /w/   *kukwega* 'to pull'

        /g/ + /w/   *tagwa* 'name'

/p/ + /y/	*-pya*	'new'
/d/ + /y/	*[lumbu] dyangu*	'my [brother]'

2.5.3. N + G combinations

A semivowel may follow a nasal (11), for the same reason it may follow a consonant. It acts like a consonant, but has the physical characteristics of a vowel.

(11)   /m/ + /w/   *kusemwa* 'to forget'

          /n/ + /y/   *kumanya* 'to know'

2.5.4. N + C + G combinations

I have found one case in my data of a glide following an NC-combination where all belong to the same syllable (12).

(12)   /n/ + /s/ + /w/   *nswa* 'termite'

2.6.   *Prosody*

Until now I have concerned myself with the segmental level of Sagala's phonology. Now I turn to the supra-segmental level, which refers to the prosodic features of a language. These features include, among others, pitch, tone, stress and intonation. Out of these, tone seems to be the most pressing concern in relation to Bantu languages, as most of these are tone languages (Kisseberth & Odden 2003: 59).

"Tone is the term used to describe the use of pitch patterns to distinguish individual words or the grammatical form of words" (Maddieson, as cited in Petzell 2008: 40-41). While minimal pairs may be used to decide whether a language is tonal or not, they are not essential for identifying a language as such. More important is the existence of contrastive lexical tonal melodies on nouns and/or verbs (Welmers 1973: 116-117). Or as McCawley says: "What is basic to the role of pitch in a tone language is not its contrastiveness but its lexicalness" (McCawley, as cited in Fromkin 1978: 3). Guthrie claims that "in Groups 30 and 40 there is neither lexical nor grammatical tone in most cases" (Guthrie 1967: 50). As far as I can tell, considering the data I have managed to collect, this claim is correct for Sagala, which belongs to the G30 group in Guthrie's referential classification of the Bantu languages.

In non-tonal languages there might, however, be stress. This refers to "portions of a word or utterance […] that are relatively prominent" (Vaux & Cooper, as cited in Petzell 2008: 41).

More air is pushed out of the lungs and this entails that the stressed sound is pronounced louder, higher and longer. While penultimate stress is common in Bantu languages (Petzell 2008: 41), Sagala has alternating stress. Although penultimate stress is the most common by far, this might be a result of Swahili influence. Stress might also be placed on the antepenult, e.g. *di'kolongo* 'hole' and *ku'hulika* 'to hear', or even before that, e.g. *'mafugamilo* 'knees'. Stress on these words is not fixed, however, as they were also pronounced with penultimate stress at a different moment. Taking these observations into consideration, there do not seem to be any obvious rules regarding stress in Sagala.

2.7.   Morphophonological processes

When two morphemes are joined, e.g. a noun prefix and a noun stem (cf. infra), certain phonological changes may occur. These changes happen in Sagala when two vowels come into contact, causing either glide formation or vowel coalescence, or in the case of nasal prefixation. Each of these processes will be explained in the following paragraphs.

2.7.1. Glide formation

When vowels occur in direct sequence, they tend to undergo either gliding or deletion. Gliding means that the front vowels /i, e/ are realised as [y] when followed by a non-identical vowel, and the back vowels /u, o/ as [w] (Hyman 2003: 48). This is a process present in Sagala. /i/ will glide to [y] in front of /a/ and /e/, whereas /u/ will glide to [w] in front of /a/ and /i/. The latter might happen with words from class 1, 3 or 15 of which the stem starts with a vowel. More on this follows later, in the section on nominal morphology. /u/ might also glide to [w] when it precedes /e/ but there is not enough evidence to back this claim. According to my observations, all these assimilations happen at the beginning of the word. In (13) you will find each instance of gliding in Sagala exemplified.

(13)   /i/ + /a/ → /ya/   i + angu   *[tate] yangu* 'my [father]'

di + angu   *[lumbu] dyangu* 'my [elder brother]'

/i/ + /e/ → /ye/   i + etu   *[tate] yetu* 'our [father]'

/u/ + /a/ → /wa/   u + angu   *[mlimo] wangu* 'my [work]'

mu + ana   *mwana* 'child'

/u/ + /i/ → /wi/   ku + inula   *kwinula* 'to lift'

ku + iguta   *kwiguta* 'to be satisfied'

2.7.2. Vowel coalescence

Vowel coalescence occurs in Sagala when the low vowel /a/ meets the high vowel /i/. This contact results in the mid vowel /e/.

(14)   /a/ + /i/ → /e/   ma + iso   *meso* 'eyes'

2.7.3. Prenasalisation

Reference was already made to prenasalisation in relation to syllable structure (cf. supra). As was said there, NC-combinations belonging to the same syllable in Sagala all consist of a prenasalised consonant. This consonant is the result of a morphophonological process whereby when a nasal meets a voiced consonant, the nasal assimilates in place to the consonant. As far as I can tell with the limited amount of data, this process is only productive in class 9/10 nouns, when the prefix *N-* meets the noun stem beginning with a voiced consonant, as shown in (15). It thus always occurs at the beginning of a word.

(15)   N + /b/ → /mb/ [mb]   N + beyu   *mbeyu* 'seed'

          N + /d/ → /nd/ [nd]   N + diya   *ndiya* 'food'

          N + /g/ → /ng/ [ŋg]   N + godi   *ngodi* 'firewood'

          N + /z/ → /nz/ [nz]   N + zila   *nzila* 'path'

2.7.4. N + voiceless stop

In the case of nasal prefixation, prenasalisation of a stop only occurs when the stop is voiced. When it is voiceless, a different phenomenon occurs which I will explain here.

Kerremans (1980) has written about a nasal followed by a voiceless consonant, and the processes that occur, both synchronically and diachronically. He treats both stops and fricatives, whereas I will limit myself to stops, since a nasal followed by a voiceless fricative, e.g. /s/, has been shown to not undergo any changes. Nasals followed by a voiceless stop, however, are submitted to the following rule: $N + C_{[-voiced]} \rightarrow N^h$. In other words, the stop is deleted and the nasal becomes aspirated. Hyman (2003: 50) calls this phenomenon de-stopping. The place of articulation of the remaining nasal will be the same as that of the deleted stop, as illustrated in the examples in (16, 17).

Aspirated nasals most commonly occur in Sagala at the beginning of class 9/10 nouns, which have an *N*-prefix (cf. infra). When this syllabic nasal meets the voiceless stop at the

beginning of the noun stem, de-stopping occurs. The data provides material to illustrate this for the voiceless velar stop /k/. Ndoweka gave *ng'huti* 'ear' with plural *makuti* 'ears'. As is clear, the plural is not formed in class 10 with prefix *N-*, but rather in class 6 with prefix *ma-*. This tells us that the noun stem for 'ear' is *-kuti*. In the case of nasal prefixation, the following occurs.

(16)   N + /k/ → /ng'h/   N + *kuti*   *ng'huti* 'ear'

When it comes to the voiceless bilabial and alveolar stops /p, t/, the data do not include sufficient material to exemplify de-stopping in case of nasal prefixation. Therefore I will base myself on comparative data and turn to the lexical Bantu reconstructions (Royal Museum for Central Africa s.d.) for help. In the examples in (17) the reconstruction of the noun is shown preceded by the noun prefix of class 9, after which the current noun in Sagala is given.

(17)   N + /p/ → /mh/   N + **pʊ̀dà*   *mhula* 'nose'

                 N + /t/ → /nh/   N + **tʊ̀ìgà*   *nhwiga* 'giraffe'

As is clear in the examples, I orthographically present the aspirated nasals as a N + h sequence. I do this in accordance to Legère (2004) and Petzell (2003, 2008) (cf. supra).

3. Nominal morphology

3.1. Nominal morphology in Bantu languages

In Bantu languages a noun fundamentally consists of the following parts:

nominal class prefix (NP) + noun stem

Apart from prefixation, suffixation also occurs in Bantu nouns, as many languages have derived nouns that are formed by this process (Katamba 2003: 103). Derivation will be treated further on in this section, in 3.3.

Nouns are subdivided into noun classes on the basis of their nominal prefix. The first one to study these classes was Brusciotto in 1659. He classified nouns on the basis of their concords. Later, Bleek was the first to reconstruct the classes of what he called 'Ancient Bantu' in 1862. He reconstructed eighteen classes and their prefixes. The next reconstruction came from the hand of Meinhof, who added five classes. Furthermore, Meeussen, Guthrie and Welmers also made reconstruction of their own. In total, twenty-four classes were reconstructed for Proto-Bantu, although no Bantu language remains which uses all twenty-four (Katamba 2003: 103-108).

Class prefixes commonly come in pairs of singular and plural. It would be a stretch to say that every singular class has its definite plural class, seeing as some nouns do not have a plural, or some nouns can form their plural in multiple classes, but there is a tendency towards it. Such pairs can also be called genders. The extent to which these genders form semantic units varies. To give an example, classes 1/2 hold human creatures and classes 3/4 typically hold nouns referring to trees and plants, but on top of that the latter also hold a disparate set of other nouns. Lastly, the noun class system regulates the concordance of noun phrase modifiers and verbs (Katamba 2003).

3.2. Nominal morphology in Sagala

In Sagala, nouns follow the fundamental structure previously mentioned. A possible addition may be a pre-prefix that is added before the NP. This will be touched upon in the paragraphs on the classes in question.

(pre-prefix (PrPr)) + NP + noun stem

Fourteen classes were found in the data. A schematic representation of all these with their corresponding noun prefixes (and pre-prefixes if they are present) is provided, accompanied by an example. Then I take a look at each class separately, giving their prefix(es), some

information on which classes they can be paired with, and the semantic content of each class. After that, I come back on the possible singular/plural class pairings.

3.2.1. Noun classes

Table 4 - Noun classes

Class	Pre-prefix	Prefix	Example
1 2		*mu- /* *m-* *wa-*	*muke* 'woman' *mlume* 'man' *wake* 'women' *walume* 'men'
3 4		*mu- /* *m-* *mi-*	*muwoko* 'hand' *mlomo* 'mouth' *milomo* 'mouths'
5 6	*di-*	*i- /* *Ø-* *ma-*	*iboma* 'wall' *tagwa* 'name' *maboma* 'walls' *matagwa* 'names'
7 8		*ki- /* *chi-* *vi- /* *fi-*	*kimage* 'knife' *chidole* 'finger' *vimage* 'knives' *fidole* 'fingers'
9 10	*i-*	*N-* *ziN-*	*nswa* 'termite' *nzagansa* 'lip' *nswa* 'termites' *zinzagansa* 'lips'
11		*lu-*	*lukolo* 'clan'
14		*u-*	*umbagi* 'spear'
15		*ku- /* *ka-*	*kuhinda* 'to shut' *kaluta* 'to go'
17		*ku-*	*kunze* 'outside' *kulwanda* 'to the river'

3.2.1.1. Class 1

Class 1 nouns have either *mu-* or *m-* as prefix. These two prefixes are in complementary distribution. *Mu-* is used with monosyllabic noun stems and stems beginning with a vowel (18), whereas *m-* is used in any other case (19). This class holds only nouns denoting human beings.

(18)   *muke*
mu-ke
NP1-woman
'woman'

mwana
mu-ana
NP1-child
'child'

(19)   *mlume*
m-lume
NP1-man
'man'

3.2.1.2. Class 2

The class 2 prefix is always *wa-* (20). If the noun stem begins with an /a/, it will coalesce with the /a/ of the prefix. There is a possibility that compensatory vowel lengthening occurs as a consequence, but as I do not have a recording of the one instance of vowel coalescence, I cannot say this with certainty. Should the stem begin with a different vowel, I assume a different consequence of vowel contact would occur, but I have not found an example of this in my data.

Class 2 forms the plural of class 1 nouns and thus includes only nouns denoting human beings as well.

(20)   *wake*
wa-ke
NP2-woman
'women'

wana
wa-ana
NP2-child
'children'

walume
wa-lume
NP2-man
'men'

3.2.1.3.   Class 3

Class 3 has prefixes *mu-* and *m-*, which are the same as those of class 1. As is the case in that class, the prefixes of class 3 are also in complementary distribution. The first occurs when the noun stem begins with a glide (21a) or a vowel (21b). It would probably also be used if the stem was monosyllabic, but no such cases were found in the data. Prefix *m-* is used in any other case (22). As far as my data shows, class 3 holds body parts, tools, and some animals.

(21)   (a)
        muwoko
        mu-woko
        NP3-hand
        'hand'

        (b)
        mwili
        mu-ili
        NP3-body
        'body'

(22)   *mlomo*
        m-lomo
        NP3-mouth
        'mouth'

        mtwango
        m-twango
        NP3-pestle
        'pestle'

        mdula
        m-dula
        NP3-hyena
        'hyena'

3.2.1.4.   Class 4

Class 4 nouns have a *mi-* prefix (23). They form the plural of class 3 nouns, even though class 3 sometimes also forms its plural in class 6, as seen in (32). I have previously talked about the gliding of /i/ into /y/ in front of vowels. I can assume that that phenomenon would occur here, but the data falls short of providing me with examples on this front. Seeing as class 4 is the regular plural class of class 3, it has the same semantic content (Katamba 2003: 115).

(23)   *milomo*
        mi-lomo

NP4-mouth
'mouths'

mitwango
mi-twango
NP4-pestle
'pestles'

midula
mi-dula
NP4-hyena
'hyenas'

3.2.1.5.   Class 5

Table 4 (cf. supra) shows that the class 5 prefixes are *i-* and *Ø-*, as shown in (24). The former is used before consonants /b, f, m, g, t, l/, while stems beginning with consonants /k, g, z, d, p, t, b, s, n, m, l/ and glides /w, y/ were given with a *Ø*-prefix. As can be seen, there is some overlap between the two. It also happened that nouns were offered with both, without a change in meaning, see (25). Overall, there seems to be a tendency to use *Ø-* as prefix.

(24)   *iboma*
        i-boma
        NP5-wall
        'wall'

        tagwa
        Ø-tagwa
        NP5-name
        'name'

(25)   (a)   (b)
        igongo   *gongo*
        i-gongo   Ø-gongo
        NP5-mountain   NP5-mountain
        'mountain'   'mountain'

Class 5 also has a pre-prefix *di-*. A pre-prefix, or augment, is another formative that precedes the prefix. Often, it is simply an additional vowel (Katamba 2003: 107). In eastern and central parts of the Bantu area, however, there are a few languages that have an augment of the CV-shape in a number of classes. This CV-augment is then formally identical with the pronominal prefixes of the corresponding classes (De Blois 1970: 93). This has led me to conclude that the *di-* with which multiple words were offered, is in fact a pre-prefix.

Very often, these words were also offered without it, or with *i-* instead, and when asked for the difference between the two forms, none was given. The function of the pre-prefix is thus as yet unclear. In Kagulu the pre-prefix has to do with definiteness of the nominal phrase and

specificity, which are in turn related to topicality (Petzell 2003: 10, 2008: 67-70). Further research would have to be done to determine whether the pre-prefix in Sagala serves a similar purpose. An example of a noun with the pre-prefix is given in (26). It is possible that instead of Ø- (26a), *i-* is used as prefix and that it has coalesced with the vowel from the pre-prefix. The noun would then be glossed as in (26b).

(26)   (a)   (b)
        dihamba   *dihamba*
        di-Ø-hamba   di-i-hamba
        PrPr5-NP5-hamba PrPr5-NP5-hamba
        'leaf'   'leaf'

Nouns with a monosyllabic stem require the use of either *i-* or *di-*. Juma and Amina consistently used *di-* (27a), whereas Ndoweka used *i-* (27b). As in the previous example, *di-* might be a coalescence of pre-prefix *di-* and prefix *i-*.

(27)   (a)   (b)
        ditwi   *itwi*
        di-Ø/i-twi   i-twi
        PrPr5-NP5-head NP5-head
        'head'   'head'

There is one noun in my data, provided in (28), that seems to belong to class 9 at first sight, seeing as it begins with an aspirated nasal. However, because its plural is formed in class 6, retaining the aspirated nasal and thus forming *mamhembe* 'horns', I assume that originally it belonged in class 9, but that it has been reanalysed as belonging in class 5.

(28)   *mhembe*
        Ø-mhembe
        NP5-horn
        'horn'

Lastly, class 5 can be used to form the augmentative of nouns. I've made this hypothesis based on the translation of 'hand' being in class 3, see (21a), whereas 'arm' is a class 5 noun, see (29). I elaborate on augmentation in 3.3.1.1.

(29)   *diwoko*
        di-woko
        NP5-hand
        'arm'

A major part of the concepts from the used wordlist belong to Class 5. However, this class does not have a lot of semantic consistency. It holds a lot of body parts, tools and instruments, some animals, and some (parts of) plants. Other than that, there are some nouns that could be identified as landscape terms and as collective nouns.

3.2.1.6.   Class 6

Class 6 nouns bear a *ma-* prefix. The plural of nouns from classes 5 and 14 are formed in this class, as seen in (30) and (31) respectively. Less regularly, it forms the plural of some nouns from classes 3, 7 and 9, see (32), (33) and (34). Other than that, there are nouns in class 6 that aren't the plural of any other noun. They are mostly mass terms and liquids (35).

(30)   *mahamba*
        ma-hamba
        NP6-leaf
        'leaves'

(31)   *mambagi*
        ma-mbagi
        NP6-spear
        'spears'

(32)   *mawoko*
        ma-woko
        NP6-hand
        'hands, arms'

(33)   *madege*
        ma-dege
        NP6-bird
        'birds'

(34)   *makuti*
        ma-kuti
        NP6-ear
        'ears'

(35)   *malovu*
        ma-lovu
        NP6-saliva
        'saliva'

When class 6 is used to form the plural of class 5, prefix-stacking might occur if the stem is monosyllabic, as in (36a). This is not a general rule, as it does not always happen (36b). However, the plural of *ditwi* 'head' might still be *maditwi* in certain syntactic contexts I have not yet encountered, thereby employing prefix-stacking as well.

(36)   (a)
        madibwe
        ma-di-Ø-bwe
        NP6-PrPr5-NP5-stone
        'stones'

(b)
matwi
ma-twi
NP6-head
'heads'

3.2.1.7. Class 7

The noun prefix of class 7 is realised differently depending on the speaker. Juma and Amina gave *ki-* as NP7 (37a), whereas Ndoweka used *chi-* (37b). As I cannot tell which is the most common in Sagala, I have included them both. It is possible, however, that the use of *ki-* is a result of the influence of Swahili. Class 7 holds mainly instruments (38), but also some animals and body parts (37). Its regular plural class is class 8, but some nouns also make their plural in class 6, as already illustrated above.

(37)   (a)
kimage
ki-mage
NP7-knife
'knife'

kidege
ki-dege
NP7-bird
'bird'

(b)
chidole
chi-dole
NP7-finger
'finger'

(38)   *kigoda*
ki-goda
NP7-chair
'chair'

3.2.1.8. Class 8

Class 8 is the regular plural class for class 7 (Katamba 2003: 115). As is the case for class 7, the NP is realised differently depending on the speaker. The realisations are *vi-* (39a) and *fi-* (39b). The use of *vi-* may again be a result of the influence of Swahili.

(39)   (a)
vimage
vi-mage

NP8-knife
'knives'

(b)
fidole
fi-dole
NP8-finger
'fingers'

3.2.1.9. Class 9

The prefix of class 9 is *N-* (40). Depending on the phoneme that follows, it can be realised in different ways. The place of articulation of the nasal will be determined by the phoneme that follows it. The different morphophonological processes that take place have been discussed in 2.7.3. and 2.7.4.

I have found two class 9 nouns which were preceded by an additional *i-*, one of which is given in (41). This is the pre-prefix. Seeing as there are only two cases, the data is too limited to define the cause for this addition.

The nouns in class 9 denote mainly animals. Body parts, tools/instruments, and some natural phenomena are found in it as well.

(40)   *nswa*
        N-swa
        NP9-termite
        'termite'

        nzagansa
        N-zagansa
        NP9-lip
        'lip'

(41)   *ing'handa*
        i-N-*gàndá
        PrPr9-NP9-house
        'house'

In (41) the Bantu lexical reconstruction (Royal Museum for Central Africa s.d.) is used for 'house', as I cannot deduce the noun stem of this noun in Sagala (cf. § 2.7.4.)

3.2.1.10. Class 10

Class 10 formally has the same prefix as class 9 (42), but it might be preceded by /zi/ (43), resulting in *ziN-* as (compound) prefix. It is the regular plural of class 9 (Katamba 2003: 115), so the semantic content is the same.

(42)　　*nswa*
　　　　N-swa
　　　　NP10-termite
　　　　'termites'

(43)　　*zinzagansa*
　　　　ziN-zagansa
　　　　NP10-nose
　　　　'noses'

3.2.1.11.　Class 11

The prefix of class 11 is *lu-* (44). I have found very few nouns of this class in my data, some of which I cannot yet be sure that they do in fact belong to it. The latter do start with *lu-*, e.g. *lufu* 'intestine', but I do not I have the plural nor a different word (i.e. an adjective, a verb, a pronoun) that is determined by it. Therefore I cannot designate these words to class 11 with certainty.

Class 11 traditionally holds long, thin entities (Katamba 2003: 115), but I have found no explicit example of this. In Swahili, class 11 has merged with class 14 (Petzell 2008: 54). This might be having an influence on Sagala, as illustrated by the fact that Ndoweka gave *lulimi* 'tongue', with a class 11 prefix, while Juma and Amina gave *ulimi*, which has a class 14 prefix (cf. infra).

(44)　　*lukolo*
　　　　lu-kolo
　　　　NP11-clan
　　　　'clan'

　　　　lulimi
　　　　lu-limi
　　　　NP11-tongue
　　　　'tongue'

3.2.1.12.　Class 12

Class 14 has *u-* as prefix (45). It appears to hold some tools, as well as abstracts and collectives. As already mentioned, if a class 14 noun has a plural, it is formed in class 6, adopting a *ma-* prefix, see example (31).

(45)　　*umbagi*
　　　　u-mbagi
　　　　NP14-spear
　　　　'spear'

umoyo
u-moyo
NP14-pity
'pity'

3.2.1.13.   Class 15

As far as my data goes, the prefix of class 15 seems to be *ku-* or *ka-* (46, 47). These would then be free allomorphs. However, Petzell (2008: 111) says that in Kagulu, the forms with *ka-*, e.g. *kaluta*, are what she refers to as the perfective, a term that indicates that "the situations [sic] is a whole "without regard to internal temporal constituency" (Comrie 1976: 2)" (Petzell 2008: 111). *Ka-* is the class 1 subject marker (SM1) used in past (PST) and perfective (PFV) (Petzell 2008: 101). *Kaluta* is to be translated with 's/he has gone' (Petzell 2008: 111). It is possible that this might also be the case in Sagala, but further research would have to be done.

(46)   *kuhinda*
        ku-hind-a
        NP15-shut-FV
        'to shut'

(47)   *kafa*
        ka-f-a
        NP15-die-FV
        'to die'

3.2.1.14.   Class 17

The class 17 prefix is *ku-* as well. There are no nouns that inherently belong to this class. *Ku-* is instead placed before the NP of other classes, as shown in (48). This process is referred to as noun-to-noun derivation (Petzell 2008: 55, 75-76) and will be further discussed in 3.3.1.3. Class 17 is one of the three locative classes in Proto-Bantu, together with class 16 and 18. The latter have not appeared in my data, but this might be because I have not focused on locations and directions. As it is, class 17 is the only locative class I have examples of and have thus been able to identify.

(48)   *kulwanda*
        ku-lu-anda
        NP17-NP11-river
        'to the river'

3.2.2. Singular/plural pairings

In the previous paragraph, I have mentioned which classes might serve as a plural. Here I will put that information together. Sagala has fourteen noun classes, which group into eight different singular/plural pairings. The most common are 1/2 , 3/4, 5/6, 7/8 and 9/10. Less common are 3/6, 7/6, 9/6 and 14/6. An overview of these findings is presented, with examples, in the following table.

Table 5 - Noun class pairings

Singular	Example	Plural	Example
Class 1	*mdala* 'old person'	**Class 2**	*wadala* 'old persons'
Class 3	*mtwango* 'pestle'	**Class 4**	*mitwango* 'pestles'
	muwoko 'hand'	**Class 6**	*mawoko* 'hands'
Class 5	*dibwe* 'stone'	**Class 6**	*mabwe* 'stones'
Class 7	*kinhu* 'thing'	**Class 8**	*vinhu* 'things'
	kidege 'bird'	**Class 6**	*madege* 'birds'
Class 9	*nswa* 'termite'	**Class 10**	*nswa* 'termites'
	ng'huti 'ear'	**Class 6**	*makuti* 'ears'
Class 11	*umbagi* 'spear'	**Class 6**	*mambagi* 'spears'

3.3.   Derivation

Derivation is a very common morphological process in Bantu languages for the formation of new words (Schadeberg 2003: 71). Sagala has both noun-to-noun and verb-to-noun derivation.

3.3.1. Noun-to-noun derivation

Throughout the previous paragraph, I have made reference to noun-to-noun derivation in the context of augmentatives and locatives. These two processes will be discussed here more elaborately.

3.3.1.1.   Augmentatives

The augmentative of a noun is formed in class 5, by replacing the noun's inherent NP by NP5. Apart from enlarging a noun's size, an augmentative might also have an affective value

(Schadeberg 2003: 83). Evidence of this has not yet been found in Sagala. The only noun that might have an augmentative connotation is *diwoko* 'arm', which is the augmentative of *muwoko* 'hand'. Seeing as I have only example (49), I cannot say with certainty that this is a productive process.

(49) *diwoko*
      di-woko
      NP5-hand
      'arm'

3.3.1.2. Other noun class shift

There exist stems which do not have an inherent class, but instead can be used in multiple ones. In Sagala I have only been able to find one of these stems, i.e. *-nhu*, which in classes 1/2 means 'person, persons', but in classes 7/8 means 'thing, things'.

3.3.1.3. Locatives

Class 17 does not have any nouns of its own, but other nouns can take its NP *ku-*. Instead of replacing the other nouns' inherent NP, *ku-* is prefixed onto the NP, resulting in a case of prefix-stacking (50). Locative derivation is a productive process.

(50) *kugongo*
      ku-Ø-gongo
      NP17-NP5-mountain
      'to the mountain'

3.3.2. Verb-to-noun derivation

Nouns might be derived from verbs by way of suffixation, though to my knowledge this is no longer a productive process in Sagala. The derived nouns have been lexicalised. Derivation has been accomplished by the use of a nominalising suffix *-i*. Morrison (2011: 171) says that in Bena classes 1/2 nouns that use the nominalising *-i* are agentive, while with other classes *-i* is used to derive results. I believe this is also the case in Sagala, as becomes clear in the examples.

Suffixation with *-i* might trigger spirantisation of the stem-final consonant. Bostoen (2008) discusses Agent Noun Spirantisation (ANS) as a particular type of Bantu Spirantisation, the latter being "the common denominator for the effect exerted by the PB [Proto-Bantu] high front and back vowel on preceding stop consonants" (Bostoen 2008: 305). ANS refers to a sound change that occurs in agent nouns derived from verbs through the use of the agentive

suffix

*-i and the nominal prefix *mʊ- (Bostoen 2008: 300).

In the examples ANS is seen to have taken place in semantically different kinds of derived nouns, as mentioned earlier. The first example (51) is an agentive noun, while the second (52) denotes a result. In both cases a voiced alveolar stop has been spirantised to a voiced alveolar fricative under the influence of the following high vowel /i/.

(51)   *-gend-a* 'to walk'   →   *m-genz-i* 'guest, traveller'

(52)   *-tund-a* 'to urinate'   →   *tunz-i* 'urine'

A second suffix that can be used to derive nouns from verbs is *o-*, which is very common in Bantu languages (Schadeberg 2003: 80-81). However, I have found only a few instances of this in my data, which are presented in (53). As /o/ is not a high vowel, it does not trigger spirantisation of the stem-final consonant.

(53)   *-fagil-a* 'to sweep'   →   *fagil-o* 'broom'

         -lim-a 'to cultivate'   →   *m-lim-o* 'work'

3.4.   *Compounding*

Some nouns seem to be the result of compounding, and have been lexicalised over time (Schadeberg 2003: 86). I have found three examples of such nouns, i.e. *mbalamwezi* 'moon', *mnyambwa* 'dog', and *kivulavumbi* 'wind'.

The first seems to be a compound of *-bad-* 'to shine,' (Royal Museum for Central Africa s.d.) and a word for 'moon'. However, as moon was given as a translation for the full word *mbalamwezi*, I cannot split the latter up, as I do not have sufficient material to back that up. The second case that appears to be the result of compounding can be split into two PB nouns, as in (54). If that is indeed the origin of that noun, the latter /a/ of *-nyama* has been deleted, and the /m/ of *-nyama* has coalesced with the NP of 'dog'. Lastly, *kivulavumbi* 'wind', can also be split into two PB words, a verb and a noun, as shown in (55). The semantic meaning of the compounds combines the semantics of each element, i.e. noun prefix, noun(s) and, where relevant, verb (cf. Morrison 2011: 175).

(54)   *mnyambwa* 'dog'   < *-nyama* 'animal' + *-bʊ̌à* 'dog'

(55)   *kivulavumbi* 'wind'   < *-púd-* 'blow' + *-gùmbɪ́* 'dust'

It should be stressed that the diachronic analyses of these compound nouns are all hypotheses.

4. Conclusion

In this paper I discussed the synchronic phonology of Sagala and presented a first overview of its noun class system. In the introduction, Sagala was situated geographically and in relation to other languages in its neighbourhood, the motivation for and the course of this research were covered and some matters regarding orthography were treated. In the section on Sagala's synchronic phonology, the phonemes were discussed, as well as the syllable structure, prosody and morphophonological processes. Sagala is a fairly typical Bantu language with a 5V system. A remarkable feature are the aspirated nasals, which are also present in a few neighbouring languages. Sagala does not have distinctive tone. It does have alternating stress.

In the section on nominal morphology, Sagala's nominal morphology was discussed by paying attention to every class individually. The noun prefixes (and pre-prefixes) and the classes' semantic content were analysed. Sagala has at least fourteen noun classes, which group into eight singular/plural pairings. It is as of yet unclear whether the three locative prefixes common in Bantu exist, due to a lack of data. The augment is present in Sagala in classes 5 and 9. Its function is still unclear.

Next to no research has previously been done on Sagala. As such, a plethora of things still remain for further research. This paper has given a first interpretation of the synchronic phonology and the nominal morphology, but because of the limited amount of data, multiple questions are still unanswered in these domains as well. With this paper I wish to contribute to the research on the languages of the Morogoro region.

5. References

Barlow, Jessica A. & Judith A. Gierut. 2002. Minimal Pair Approaches to Phonological Remediation. *Seminars in Speech and Language* 23/1, 57-67.

Batibo, Herman M. 2009. Language Documentation as a Strategy for the Empowerment of the Minority Languages of Africa. In Masangu Matondo, Fiona Mc Laughlin & Eric Potsdam (eds.) *Selected Proceedings of the 38th Annual Conference on African Linguistics*, 193-203. Somerville, MA: Cascadilla Proceedings Project.

Bostoen, Koen. 2008. Bantu Spirantization: Morphologization, Lexicalization and Historical Classification. *Diachronica* 25/3, 299-356.

Bowern, Claire. 2008. *Linguistic fieldwork: A practical guide.* Hampshire; New York: Palgrave Macmillan.

De Blois, K.F. 1970. The Augment in the Bantu Languages. *Africana Linguistica* 4, 85-165.

Fromkin, Victoria A. 1978. *Tone: A linguistic survey*. London; New York: Academic Press.

Gonzales, Rhonda Marie. 2002. *Continuity and change: thought, belief, and practice in the history of the Ruvu peoples of Central East Tanzania, c. 200 B.C. to A.D. 1800.* PhD. dissertation, University of California.

Guthrie, Malcolm. 1967 [1948]. *The classification of the Bantu languages*. London: Dawsons of Pall Mall for the International African Institute.

Hayes, Bruce. 2009. *Introductory Phonology*. West Sussex: Wiley-Blackwell.

Himmelmann, Nikolaus P. 1998. Documentary Linguistics. *Linguistics* 36, 161-195.

Hyman, Larry M. 2003. Segmental phonology. In Derek Nurse & Gérard Philippson (eds.), 42-58.

Katamba, Francis. 2003. Bantu nominal morphology. In Derek Nurse & Gérard Philippson (eds.), 103-120.

Kerremans, R. 1980. Nasale Suivie de Consonne Sourde en Proto-Bantou. *Africana Linguistica* 8, 159-198.

Kisseberth, Charles & David Odden. 2003. Tone. In Derek Nurse & Gérard Philippson (eds.), 59-70.

Lass, Roger. 1984. *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Legère, Karsten, S. Maganga & P. Mkwan'hembo. 2004. Vidunda People and their Plant Names. *Africa & Asia* 4, 115-141.

Legère, Karsten. 2006. Language Endangerment in Tanzania: Identifying and Maintaining Endangered Languages. *South African Journal of African Languages* 26/3, 99-112.

Legère, Karsten. 2007. Vidunda (G38) as an Endangered Language? In Doris L. Payne and Jaime Peña (eds.) *Selected Proceedings of the 37th Annual Conference on African Linguistics*, 43-54. Somerville, MA: Cascadilla Proceedings Project, 2007.

Legère, Karsten. 2009. Plant Names in the Tanzanian Bantu Language Vidunda: Structure and (Some) Etymology. In Masangu Matondo, Fiona Mc Laughlin & Eric Potsdam (eds.) *Selected Proceedings of the 38th Annual Conference on African Linguistics*, 217-228. Somerville, MA: Cascadilla Proceedings Project.

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 (25.03.2017).

Meeuwis, Michael & Hilde Gunnink. 2014. *Inleiding tot de Afrikaans taalkunde: band 1*. Gent: Academia Press.

Meinhof, Carl & Daniel Jones. 1928. Principles of Practical Orthography for African Languages. *Africa: Journal of the International African Institute* 1/2, 228-239.

Morrison, Michelle Elizabeth. 2011. *A reference grammar of Bena.* PhD. dissertation, Rice University.

Muzale, Henry R.T. & Josephat M. Rugemalira. 2008. Researching and Documenting the Languages of Tanzania. *Language Documentation and Conservation* 2/1, 68-108.

Nurse, Derek & Gérard Philippson. 1980. The Bantu languages of East Africa: A lexicostatistical survey. In Edgar C. Polomé & C.P. Hill (eds.), 26-67.

Nurse, Derek & Gérard Philippson. 2003. *The Bantu languages*. London; New York: Routledge.

Petzell, Malin. 2003. What is the Function of the Pre-Prefix in Kagulu? *langue.doc.* https://gupea.ub.gu.se/bitstream/2077/24/5/Petzell_M_2003.pdf (1.02.2017).

Petzell, Malin. 2008. *The Kagulu grammar of Tanzania: Grammar, texts and vocabulary*. Cologne: Rüdiger Köppe Verlag.

Petzell, Malin & Harald Hammarström. 2013. Grammatical and Lexical Comparison of the Greater Ruvu Bantu Languages. *Nordic Journal of African Studies* 22/3, 129-157.

Polomé, Edgar C. 1980. The languages of Tanzania. In Edgar C. Polomé & C.P. Hill (eds.), 3-25.

Polomé, Edgar C. & C.P. Hill. 1980. *Language in Tanzania*. Oxford: Oxford University Press for the International African Institute.

Royal Museum for Central Africa. s.d. [Bantu Lexical Reconstructions 3]. http://www.africamuseum.be/collections/browsecollections/humansciences/blr (6.05.2017)

Schadeberg, Thilo C. 2003. Derivation. In Derek Nurse & Gérard Philippson (eds.), 71-89.

Schadeberg, Thilo C. 2009. Loanwords in Swahili. In Martin Haspelmath and Uri Tadmor (eds.) *Loanwords in the World's Languages: a Comparative Handbook*, 76-102. Berlin: De Gruyter Mouton.

Seifart, Frank. 2006. Orthography Development. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.) *Essentials of Language Documentation*, 275-300. Berlin; New York: Mouton de Gruyter.

Soto-Faraco, Salvador & Núria Sebastián-Gallés. 2001. Segmental and Suprasegmental Mismatch in Lexical Access. *Journal of Memory and Language* 45, 412-432.

UNESCO Ad Hoc Expert Group on Endangered Languages. 2003. Language vitality and endangerment. Document submitted to *The International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages*, Paris, 10-12 March 2003.

Welmers, William E. 1973. *African language structures*. Berkeley: University of California Press.

Zhang, Jisheng. 2006. *The phonology of Shaoxing Chinese.* PhD. dissertation, University of Leiden.

Appendix A: Wordlists

The words are ordered alphabetically according to the English translation, in order to make comparing with the second wordlist easier. Other than English, the words are offered in Swahili and Sagala.

Class 5 nouns are presented as they were offered by the speakers. This means that if a noun was only offered either with PrPr5 *di-* or NP5 *i-*, it is presented that way, e.g. *dihamba* 'leaf', and *ibahu* 'year'. If it was offered with NP5 Ø- as well as with either of the previous ones, *di-* and *i-* are put between brackets, e.g. *(di)fagilo* 'knee' and *(i)gongo* 'mountain'. If the class 5 noun was offered with all three options, it is presented as follows, e.g. *(di)/(i)kolongo* 'hole'. Verbs are presented with NP15 *ku-* or *ka-*, sometimes both, depending on which was used during the elicitation sessions.

If the plural of a noun is available, it is given immediately after the singular, separated by a comma. If there are multiple translations for a word, they are presented separated by '/'. Where known, the noun classes are provided, both for singular (NCS) and plural (NCP). A question mark is put in case of uncertainty.

1. Wordlist Juma and Amina

English	Swahili	Sagala	NCS	NCP	Part of speech
abdomen	*tumbo*	*inda*	?		n.
ancestral spirit	*mzimu*	*mzimu, mizimu*	3	4	n.
anger	*hasira*	*maya*	?		n.
animal	*mnyama*	*mnyama, wanyama*	1	2	n.
ankle	*fundo la mguu*	*kibwengu*	7		n.
arm	*mkono*	*diwoko, mawoko*	5	6	n.
armpit	*kwapa*	*ng'hwapa, ng'hwapa*	9	10	n.
arrow	*mshale*	*msile, misile*	3	4	n.
ashes	*majivu*	*madivu*		6	n.
axe	*shoka*	*nhemo, nhemo*	9	10	n.
back (of body)	*ngongo*	*digongo*	5		n.
bad	*-baya*	*-ha*			adj.
barren woman	*mwanamke tasa*	*muke mgumba, wake wagumba*	1	2	n.
basket	*kikapu*	*ndoto, mandoto*	5	6	n.
beard	*ndevu*	*ndevu*	9		n.

bed	*kitanda*	*sazi, masazi*	5	6	n.
big	*-kubwa*	*-kulu*			adj.
bird	*ndege*	*kidege, madege*	7	6	n.
bitter	*chungu*	*usungu*	14		n.
black	*-eusi*	*-titu*			adj.
blood	*damu*	*sakame*	?		n.
body	*mwili*	*mwili*	3		n.
bone	*mfupa*	*guha, maguha*	5	6	n.
boundary	*mpaka*	*mbaka*	?		n.
bow	*upinde*	*upinde*	14		n.
boy	*mvulana*	*mbwanga, wabwanga / mgosi*	1	2	n.
brain	*ubongo*	*ubongo*	14		n.
branch	*tawi*	*ditawi, matawi*	5	6	n.
breast(s) female	*(ma)titi*	*tombo, matombo*	5	6	n.
broom	*ufagio*	*(di)fagilo, mafagilo*	5	6	n.
buffalo	*nyati, mbogo*	*(i)palati*	?		n.
bundle of firewood	*mzigo wa kuni*	*dizigo la ngodi*	5		n.
buttocks	*tako*	*dako, madako*	5		n.
canoe	*mtumbwi*	*mtumbwi, mitumbwi*	3	4	n.
cattle	*ng'ombe*	*ng'ombe*	9		n.
chair	*kiti*	*kigoda, magoda*	7	6	n.
charcoal	*mkaa*	*mkala*	3		n.
cheek	*shavu*	*kutu, makutu / disavu, masavu*	5	6	n.
chest	*kifua*	*kipehema*	7		n.
chicken	*kuku*	*ng'huku, ng'huku*	9	10	n.
child	*mwana, mtoto*	*mwana, wana*	1	2	n.
chin	*kidevu*	*kidevu*	7		n.
clan	*ukoo*	*dari / lukolo*	? / 11		n.
clothing	*nguo*	*suke / mwenda*	?		n.
cloud	*uwingu*	*wingu, mawingu*	5	6	n.
cold	*baridi*	*-beho*			adj.
conversation	*mazungumzo*	*mbuli ya kulonga*	9		n.
cooking stones	*mafiga*	*mafiga*		6	n.
corpse	*maiti*	*maiti*	?		n.
crocodile	*mamba*	*mamba, mamba*	9	10	n.
darkness	*giza*	*finzi*	?		n.
daytime	*mchana*	*nemisi*	?		n.
death	*kifo*	*difa*	5		n.
desire	*tamaa*	*ngiri*	9		n.
dew	*umande*	*umande*	14		n.
dirty	*chafu*	*usafu*	14		n.
dog	*mbwa*	*mnyambwa, wanyambwa*	1	2	n.

door	*mlango*	*ibakwa, mabakwa*	5	6	n.
drum	*ngoma*	*ngoma*	9		n.
dry	*-kavu*	*-kavu*			adj.
dust	*vumbi*	*ng'hundi*	9		n.
ear	*sikio*	*guti, maguti*	5	6	n.
egg	*yai*	*ganga, maganga*	5	6	n.
eight	*nane*	*manane*			num.
eighty	*themanini*	*makumi manane*			num.
elder brother	*kaka*	*lumbu*	?		n.
elder sister	*dada*	*lumbu*	?		n.
elephant	*tembo / ndovu*	*nhembo, nhembo*	9	10	n.
eleven	*kumi na moja*	*makumi dimwe*			num.
evening	*jioni*	*nyamihe*	?		n.
eye	*jicho*	*ziso, meso*	5	6	n.
face	*uso*	*kihanga, mahanga*	7	6	n.
far	*mbali*	*kutali*			adv.
fat (person)	*-nene*	*-neneha*			adj.
father	*baba*	*tate*	?		n.
feather	*unyoya wa ndege*	*unyoya, manyoya*	14	6	n.
fever	*homa*	*homa*	?		n.
field	*shamba*	*mgunda, migunda*	3	4	n.
fifteen	*kumi na tano*	*makumi ya ihano*			num.
fifty	*hamsini*	*makumi mahano*			num.
finger	*kidole*	*kidole*	7		n.
fingernail	*ukucha*	*kombe, makombe*	5	6	n.
fire	*moto*	*mlilo*	3		n.
fireplace	*jiko*	*diziko, maziko*	5	6	n.
firewood	*kuni*	*ngodi*	9		n.
fish	*samaki / nswi*	*somba*	?		n.
fishhook	*ndoano*	*ndoano, ndoano*	9	10	n.
five	*tano*	*ihano*			num.
flesh	*nyama*	*nyama*	9		n.
flower	*ua*	*luwa, maluwa*	5	6	n.
food	*chakula*	*ndiya*	9		n.
foot	*mguu*	*mgulu, magulu*	3	6	n.
fourty	*arobaini*	*makumi maane*			num.
fruit	*tunda*	*tunda, matunda*	5	6	n.
giraffe	*twiga*	*nhwiga*	9		n.
girl	*msichana*	*kigoli, vigoli*	7	8	n.
goat	*mbuzi*	*mhene, mhene*	9	10	n.
good	*-ema*	*-noga*			adj.
gourd	*buyu*	*mung'unya*	?		n.

grass	nyasi	nyasi	9		n.
guest	mgeni	mgenzi, wagenzi	1	2	n.
guinea fowl	kanga	ng'hanga, ng'hanga	9	10	n.
hair (of head)	nywele	mvili	9		n.
hammer	nyundo	imundo	?		n.
handle	mpini	muhini	3		n.
hard	-gumu	-kugutu			adj.
head	kichwa	ditwi, matwi	5	6	n.
heart	moyo				n.
heat	joto	ivuke	?		n.
heel	kisigino				n.
hippopotamus	kiboko	tomondo	?		n.
his father	baba yake	tate yake			n. + pron.
his mother	mama yake	mayi yake			n. + pron.
hole	shimo	ikolongo (di-), makolongo	5	6	n.
honey	asali	uki	14		n.
honey bee	nyuki	nzuki, nzuki	9	10	n.
horn	pembe	mhembe, mamhembe	5	6	n.
house	nyumba	(i)ng'handa	9		n.
house fly	nzi	ng'honzi, ng'honzi	9	10	n.
hunger	njaa	nzala	9		n.
hunter	mwindaji	mlumba, walumba	1	2	n.
husband	mume	mlume, walume	1	2	n.
hyena	fisi	mdula, midula	3	4	n.
infant	mtoto mchanga	mwana msanga, watoto wasanga	1	2	n.
intestine	utumbo	lufu	?		n.
knee	goti	(i)fugamilo, mafugamilo	5	6	n.
knife	kisu	kimage, vimage	7	8	n.
knot	fundo	ifundo, mafundo	5	6	n.
lame	kiwete	kimbete	7		n.
lamp	taa	koroboi	?		n.
land	nchi (ardhi)	isi	?		n.
language	lugha	lugaki	?		n.
leaf	jani	dihamba, mahamba	5	6	n.
left hand	mkono wa kushoto	muwoko wa kushoto, mawoko	3	6	n.
leopard	chui	dume, dume	9	10	n.
lie	uwongo	kubela	15		v.
light	nuru / mwangaza	mwanga	?		n.
light	-epesi	-pelu			adj.
lightning	radi	radu	?		n.
lion	simba	makangila	?		n.
lip	mdomo	nzagansa, zinzagansa	9	10	n.

liver	ini	itoga	?		n.
load	mzigo	dizigo, mazigo	5	6	n.
long	-refu	-tali			adj.
louse (of body)	chawa	mmani, mmani	9	10	n.
lungs	mapafu	wengu, mawengu	5	6	n.
man	mwanaume	mlume, walume	1	2	n.
marriage	ndoa	posa	?		n.
mat	jamvi / mkeka	kilago, vilago	7	8	n.
matchet	panga	mundu, zimundu / ipanga	9 / 5	10/ ?	n.
meat	nyama	nyama	9		n.
medicine	dawa	miti	?		n.
medicine man	mganga	mganga, waganga	1	2	n.
milk	maziwa	mele	?		n.
monkey	tumbili	ngedele, ngedele	9		n.
moon	mwezi	mbalamwezi	?		n.
morning	asubuhi	hegulo	?		n.
mortar	kinu	(di)/(i)tuli, matuli	5	6	n.
mosquito	mbu	nzuguni, nzuguni	9	10	n.
mother	mama	mayi	?		n.
mountain	mlima	(i)gongo, magongo	5	6	n.
mouth	kinywa	mlomo, milomo	3	4	n.
mud	tope	nhope	9		n.
mushroom	uyoga	uyoga	14		n.
my father	baba yangu	tate yangu			n. + pron.
my mother	mama yangu	mayi yangu			n. + pron.
name	jina	tagwa, matagwa	5	6	n.
narrow	-embamba	-sisili			adj.
navel	kitovu	duwo, maduwo	5	6	n.
near	karibu	karibu			adv.
neck	shingo	singo	?		n.
new	-pya	-pya			adj.
night	usiku	kilo	?		n.
nine	tisa	ikenda			num.
nineteen	kumi na tisa	makumi kenda			num.
ninety	tisini	makumi ikenda			num.
nose	pua	mhula, zimhula	9	10	n.
nostrils	mashimo ya pua				n.
oil	mafuta	mafuta		6	n.
old	-kuukuu	-sakala			adj.
old person	mzee	mdala, wadala	1	2	n.
one	moja	imwe			num.
one hundred	mia	igaana			num.
other	-ingine	-hage			adj.

our father	*baba yetu*	*tate yetu*			n. + pron.
pain (of body)	*maumivu*	*-bulagika*			v.
palm (of hand)	*kiganja*	*ganza, maganza*	5	6	n.
partridge	*kwale*	*ng'hwale, ng'hwale*	9	10	n.
path	*njia*	*nzila, nzila*	9	10	n.
person	*mtu*	*munhu*	1		n.
pestle	*mchi*	*mtwango, mitwango*	3	4	n.
pig	*nguruwe*	*ngubi, ngubi*	9	10	n.
pity	*huruma*	*umoyo*	14		n.
place	*mahali*	*ahano*	?		n.
pot	*chungu*	*nyungu, nyungu / nhongo, zinhogo*	9	10	n.
proud	*enye majivuno*	*nyodo*	?		n.
python	*chatu*	*mgowe, migowe*	3	4	n.
quarrel	*ugomvi*	*vagi*	?		n.
rain	*mvua*	*mvula*	9		n.
rat	*panya*	*ngule, mangule*	5	6	n.
red	*-ekundu*	*-dunhu*			adj.
rib	*ubavu*	*paga, mapaga*	5	6	n.
right hand	*mkono wa kulia*	*muwoko wa kulia, mawoko*	3	6	n.
river	*mto*	*lwanda*	11		n.
root	*mzizi*	*mzizi, mizizi*	3	4	n.
rope	*kamba*	*nhamba, nhamba*	9	10	n.
saliva	*mate*	*malovu*		6	n.
salt	*chumvi*	*munyu*	?		n.
sand	*mchanga*	*msanga*	3		n.
scar	*kovu*	*bangu, mabangu*	5	6	n.
seed	*mbegu*	*mbeyu, mbeyu*	9	10	n.
seven	*saba*	*vihaba*			n.
seventeen	*kumi na saba*	*makumi ya vihaba*			num.
seventy	*sabini*	*makumi vihaba*			num.
shadow	*kivuli*	*lubweho*	?		n.
shame	*aibu*	*kinyala*	?		n.
sheep	*kondoo*	*ng'holo, ng'holo*	9	10	n.
short	*-fupi*	*-guhe*			adj.
shoulder	*bega*	*yega, mayega*	5	6	n.
six	*sita*	*izita*			num.
sixteen	*kumi na sita*	*makumi ya izita*			num.
sixty	*sitini*	*makumi izita*			num.
skin	*ngozi*	*ng'hingo*	9		n.
sleep	*usingizi*	*honga*	?		n.
small	*-dogo*	*-dodo*			adj.
smoke	*moshi*	*mosi*	?		n.

snake	nyoka	(di)zoka, mazoka	5	6	n.
soil (clay)	udongo	idongo	5		n.
soldier ant	chungu	nyakongo	9		n.
song	wimbo	nyimbo, nyimbo	9	10	n.
spear	mkuki	umbagi, mambagi	14	6	n.
star	nyota	nyota, nyota	9	10	n.
stick	fimbo	ng'home, ng'home	9	10	n.
stone	jiwe	dibwe, madibwe	5	6	n.
sugar-cane	muwa	mguwa	3		n.
sun	jua	dizua	5		n.
sweat	jasho	divuke	5		n.
sweet	tamu	kenoga / kunoga	15		v.
tail	mkia	mkila / dikila, mikila	3 / 5	4	n.
tear(s)	machozi	mahozi		6	n.
ten	kumi	ikumi			num.
termite	mchwa	nswa, nswa	9	10	n.
their father	baba yao	tate yao			n. + pron.
their mother	mama yao	mayi yao			n. + pron.
thin	-embamba	-sisili			adj.
thing	kitu	kinhu, vinhu	7	8	n.
thirst	kiu	ng'halu	9		n.
thirty	thelathini	makumi madatu			num.
thorn	mwiba	imuwa, mamuwa	5	6	n.
three	tatu	idatu			num.
thunder	ngurumo	mlindimo / dangubuma	?		n.
to answer	kujibu	kagalilwa	15		v.
to arrive	kufika				v.
to ask (question)	kuuliza	kuuza	15		v.
to ask for	kuomba	kulanza	15		v.
to be	kuwa	sambi	?		v.
to be angry	kasirika	kena maya	15		v.
to be astonished	kushangaa	mbelula	?		v.
to be born	kuzaliwa	kaleligwa	15		v.
to be full	kujaa	kumema	15		v.
to be married	kuolewa	kutoligwa / katoligwa	15		v.
to be rotten	kuoza	-doda			adj.
to be satisfied	kushiba	kwiguta	15		v.
to be tired	kuchoka	kusoka	15		v.
to beat	kupiga kofi	kutowa diganja	15		v.
to begin	kuanza	kukonga	15		v.
to belch	kubeua	kubeua	15		v.
to bend	kupinda	kuzinga / kezinga	15		v.
to bite	kuuma	kuluma	15		v.

to blow (wind)	*kuvuma*	*kulindima*	15		v.
to boil (intr.)	*kuchemka*	*kuhemsa*	15		v.
to break (tr.)	*kuvunja*	*kubena*	15		v.
to breathe	*kupumua*	*kuhumula*	15		v.
to bring	*kuleta*	*kugala / kagala*	15		v.
to build	*kujenga*	*kuzenga*	15		v.
to burn up	*kuchoma*	*kuhoma*	15		v.
to bury	*kuzika*	*kuzika*	15		v.
to buy	*kununua*	*kugula*	15		v.
to call	*kuita*	*kukema*	15		v.
to carry water	*kuchukua maji*	*kusola mezi*	15		v.
to chase away	*kufukuza*	*kuwinga*	15		v.
to come	*kuja*	*kwiza*	15		v.
to come from	*kutoka*	*kulawa (hoki)*	15		v.
to cool down	*kupoa*	*kuhola*	15		v.
to cough	*kukohoa*	*ng'hololo*	?		v.
to count	*kuhesabu*	*kuhesabu*	15		v.
to cover (a pot)	*kufunika*	*kugubika*	15		v.
to cry	*kulia*	*kulila*	15		v.
to cultivate	*kulima*	*kulima*	15		v.
to cure	*kuponya ugonjwa*	*kuhonya*	15		v.
to curse	*kulaani*	*kusua*	15		v.
to cut	*kukata*	*kutema*	15		v.
to deceive	*kudanganya*	*kudana*	15		v.
to decorate	*kupamba*	*kuhamba*	15		v.
to die	*kufa*	*kafa*	15		v.
to dig (hole)	*kuchimba*	*kuhimba*	15		v.
to divide	*kugawanya*	*kugawila*	15		v.
to do	*kufanya*	*kutenda*	15		v.
to draw water	*kuteka maji*	*kunega mezi*	15		v.
to dream	*kuota*	*kulota*	15		v.
to dress (tr.)	*kuvika*	*kuvika*	15		v.
to drink	*kunywa*	*kunyuwa*	15		v.
to eat	*kula*	*kudiya*	15		v.
to enter	*kuingia*	*kuingila*	15		v.
to explain	*kueleza*	*kulonga mbuli*	15		v.
to extinguish	*kuzima*	*kuzima*	15		v.
to fall	*kuanguka*	*kugwa / kagwa*	15		v.
to fall ill	*kuugua*	*kutamwa*	15		v.
to fight	*kupigana*	*kutowana*	15		v.
to fill	*kujaza*	*kumema*	15		v.
to finish	*kumaliza*	*kumala*	15		v.

to finish (tr.)	*kumaliza*	*kumala*	15		v.
to fish	*kuvua samaki*	*kulowa somba*	15		v.
to fly	*kuruka*	*kuzuma*	15		v.
to forget	*kusahau*	*kusemwa*	15		v.
to fry	*kukaanga*	*kukalanga*	15		v.
to gather (fruits)	*kuchuma matunda*	*kubawa matunda*	15		v.
to get	*kupata*	*keng'higwa*	15		v.
to get drunk	*kulewa*	*kulewa*	15		v.
to get hurt	*kuumia*	*kubulaga*	15		v.
to give	*kupa*	*kaheleza*	15		v.
to give birth	*kuzaa*	*kulela*	15		v.
to go	*kuenda*	*kaluta / kuluta*	15		v.
to go away	*kuondoka*	*kauka*	15		v.
to grind (grain)	*kusaga*	*kusaga*	15		v.
to hang up	*kutundika*	*nhungika*	?		v.
to harvest	*kuvuna*	*kugola*	15		v.
to hear	*kusikia*	*kuhulika*	15		v.
to help	*kusaidia*	*kutaza*	15		v.
to hide (something)	*kuficha*	*kufisa*	15		v.
to hunt	*kuwinda*	*kusaka*	15		v.
to insult	*kutukana*	*kuliga*	15		v.
to jump	*kuruka*	*kuzuma*	15		v.
to kill	*kuua*	*kukoma*	15		v.
to kindle (fire)	*washa*	*kupemba (mlilo)*	15		v.
to know	*kujua*	*kumanya*	15		v.
to laugh	*kucheka*	*kuseka*	15		v.
to lay down	*kuweka*	*kuwika*	15		v.
to learn	*kujifunza*	*kufunda*	15		v.
to lift	*kunyanyua*	*kwinula*	15		v.
to look after a child	*kulea*	*kulela*	15		v.
to marry	*kuoa*	*kutola*	15		v.
to measure	*kupima*	*mmaha*	?		n.
to milk	*kukamua*	*kukama mele*	15		v.
to mix	*kuchanganya*	*kuhanza*	15		v.
to mould	*kufinyanga*	*kuumba*	15		v.
to open	*kufungua*	*kudigula*	15		v.
to pass	*kupita*	*kupuluta*	15		v.
to pay	*kulipa*	*kuliha*	15		v.
to plait	*kusuka nywele*	*kusuka mvili*	15		v.
to plant	*kupanda (mbegu)*	*kuhanda (mbeyu)*	15		v.
to play	*kucheza*	*kuvina*	15		v.

to pound	kutwanga	kutwanga	15		v.
to pull	kuvuta	kukwega	15		v.
to push	kusukuma	kubimbiligisa	15		v.
to put into	kutia	kugela	15		v.
to quarrel	kugombana	kugobola vagi	15		v.
to rain	kunyesha	kutonya	15		v.
to rear	kufuga	kufuga	15		v.
to refuse	kukataa	kulema / kalema	15		v.
to rest	kupumzika	kubuhila	15		v.
to return	kurudi	kuuya	15		v.
to run away	kukimbia	kubilima	15		v.
to see	kuona	kulola	15		v.
to seize	kushika	kukola	15		v.
to sell	kuuza	kuguza	15		v.
to send	kutuma	kakihilika	15		v.
to sew	kushona	kuhona	15		v.
to shake (tr.)	kutikisa	kusingisa	15		v.
to shine	kung'aa	kung'ala	15		v.
to shout	sema kwa sauti kubwa	kulonga kwa lwangi	15		v.
to show	kuonyesha	kulanga	15		v.
to shut	kufunga	kuhinda	15		v.
to sing	kuimba	kwimba	15		v.
to slap	kupiga kofi	kutowa diganja	15		v.
to sleep	kulala	kugona	15		v.
to smash	kuponda	kuhonda	15		v.
to sneeze	kupiga chafya	kulava nyasa	15		v.
to speak	kuongea	kulonga	15		v.
to split (firewood)	kuchanja (kuni)	kutema	15		v.
to sprout	kuchipua	kudedewala	15		v.
to stand	kusimama	kuwima	15		v.
to steal	kuiba	kuhiza	15		v.
to stink	kunuka	kunung'ha	15		v.
to swallow	kumeza	kumela	15		v.
to sweep	kufagia	kufagila	15		v.
to take	kuchukua	kusola	15		v.
to take a bath	kuoga	kwiyoga	15		v.
to taste (food)	kuonja	kugeza	15		v.
to teach	kufundisha	kufundigwa	15		v.
to tear	kupasua	kutula	15		v.
to tell	kuambia	kemba / kulongigwa	15		v.
to throw	kutupa	kutaga	15		v.
to try	kujaribu	kugeza	15		v.

to twist (rope)	*sokota (kamba)*	*kusogota*	15		v.
to urinate	*kukojoa*	*kutunda*	15		v.
to vomit	*kutapika*	*kudeka*	15		v.
to wait	*kungoja*	*kubeta*	15		v.
to walk	*kutembea*	*kugenda*	15		v.
to wash (hands)	*kunawa mikono*	*kuhovuga mawoko*	15		v.
to wash clothes	*kufua nguo*	*kuhovuga suke*	15		v.
to wear	*kuvaa*	*kuvala*	15		v.
to work	*kufanya kazi*	*kukola milimo*	15		v.
to wound	*kuumiza*	*kambuaga*	15		v.
today	*leo*	*diyelo*			adv.
tomorrow	*kesho*	*mitondo*			adv.
tongue	*ulimi*	*ulimi*	14		n.
tooth	*jino*	*zino, mazino*	5	6	n.
trap	*mtego*	*mnhego, minhego*	3	4	n.
tree	*mti*	*(di)biki, mabiki*	5	6	n.
twenty	*ishirini*	*makumi meeli*			num.
twins	*mapacha*	*mapinga*	6		n.
unripe	*-bichi*	*-bisi*			adj.
urine	*mkojo*	*tunzi*	?		n.
valley	*bonde*	*(di)/(i)lolo*	5		n.
voice	*sauti*	*lwangi*	?		n.
wall	*ukuta*	*iboma, maboma*	5	6	n.
water	*maji*	*mezi*	?		n.
white	*-eupe*	*-zelu*			adj.
wide	*-pana*	*-gazi*			adj.
wife	*mke*	*muke, wake*	1		n.
wind	*upepo*	*kivulavumbi*	7		n.
wing	*ubawa*	*kipapatilo, mapapatilo*	7	6	n.
woman	*mwanamke*	*muke, wake*	1		n.
work	*kazi*	*mlimo*	3		n.
year	*mwaka*	*ibahu, mabahu*	5	6	n.
yesterday	*jana*	*ijana*			adv.
your (pl.) father	*baba yenu*	*tate yenu*			n. + pron.
your (pl.) mother	*mama yenu*	*mayi yenu*			n. + pron.
your father	*baba yako*	*tate yako*			n. + pron.
your mother	*mama yako*	*mayi yako*			n. + pron.

2. Wordlist Ndoweka

English	Swahili	Sagala	NCS	NCP	Part of speech
abdomen	*tumbo*	*ida, mada*	5	6	n.
ankle	*fundo la mguu*	*chisukusuku / ng'hoko*	7 / 9		n.
arm	*mkono*	*iwoko*	5		n.
armpit	*kwapa*	*ng'hwawa*	9		n.
ashes	*majivu*	*machifu*		6	n.
axe	*shoka*	*nhemo*	9		n.
back (of body)	*ngongo*	*mgongo*	?		n.
basket	*kikapu*	*sege*	?		n.
bed	*kitanda*	*isasi*	5		n.
blood	*damu*	*sakame / usaha / ufila*	?		n.
body	*mwili*	*mtufi*	3 ?		n.
bone	*mfupa*	*(i)guha, maguha*	5	6	n.
boy	*mvulana*	*mwanike / msongolo*	1 ?		n.
brain	*ubongo*	*chawongo*	7		n.
breast(s) female	*(ma)titi*	*tombo, matombo*	5	6	n.
broom	*ufagio*	*fagio*	?		n.
bundle of firewood	*mzigo wa kuni*	*msigo wa ngodi*	3		n.
buttocks	*tako*	*dako, madako*	5	6	n.
chair	*kiti*	*chigoda*	7		n.
charcoal	*mkaa*	*mkala*	3		n.
cheek	*shavu*	*jeje, majeje*	5	6	n.
child	*mwana, mtoto*	*mwana*	1		n.
chin	*kidevu*	*chidevu*	7		n.
clan	*ukoo*	*kuungugo / mtala*	?		n.
cooking stones	*mafiga*	*figa, mafiga*	5	6	n.
ear	*sikio*	*ng'huti, makuti*	9	6	n.
elbow	*kivi*	*chisukusuku*	7		n.
elder brother	*kaka*	*kaka*	?		n.
elder sister	*dada*	*lumbu*	?		n.
elephant	*tembo / ndovu*	*nhembo*	9		n.
eye	*jicho*	*igiso, magiso*	5		n.
face	*uso*	*chihanga*	7		n.
father	*baba*	*tate*	?		n.
finger	*kidole*	*chidole, fidole*	7	8	n.
fingernail	*ukucha*	*kombe, makombe*	5	6	n.
fire	*moto*	*moto*	?		n.
fireplace	*jiko*	*chiko*	?		n.
firewood	*kuni*	*ngodi*	9		n.

flesh	nyama	nyama	9		n.
foot	mguu	lwatilo	?		n.
force	nguvu	ludole	?		n.
gourd	buyu	buyu	?		n.
guest	mgeni	mgeni / kasamila / yeja yeja	1 / ?		n.
hair (of head)	nywele	njwili	9		n.
hammer	nyundo	nyundo, nyundo	9	10	n.
handle	mpini	muhini	?		n.
head	kichwa	ditwi / itwi, matwi	5	6	n.
heart	moyo	moyo	?		n.
heel	kisigino	chisigino	7		n.
his father	baba yake	tate yake	?		n. + pron.
house	nyumba	kaya	?		n.
husband	mume	mgosi	1		n.
infant	mtoto mchanga	chana chilele / mwana mlele	7 / 1		n.
intestine	utumbo	(u)tumbu	?		n.
kidney	figo	figo	?		n.
knee	goti	(i)findi, mafindi	5	6	n.
knife	kisu	(u)mage, mage	?		n.
knot	fundo	fundo	?		n.
left hand	mkono wa kushoto	diganja dya kumoso	5		n.
leg	mguu	mgulu	3		n.
lie	uwongo	udasi	14		v.
liver	ini	toga, matoga	5	6	n.
lungs	mapafu	mapafu		6	n.
man	mwanaume	mgosi	1		n.
marriage	ndoa	lusona	11		n.
mat	jamvi / mkeka	mkeka / jamvi	?		n.
matchet	panga	panga	?		n.
mortar	kinu	tuli, matuli	5	6	n.
mother	mama	mayi	?		n.
mouth	kinywa	mlomo	3		n.
my mother	mama yangu	mayi yangu	?		n. + pron.
name	jina	tagwa	?		n.
navel	kitovu	chiduwo / (i)duwo	7 / 5		n.
neck	shingo	singo, singo	9	10	n.
nose	pua	mhula	9		n.
nostrils	mashimo ya pua	makobombo ga / ya mhula		6	n.
old person	mzee	mdala	1		n.
our father	baba yetu	tate yetu	?		n. + pron.
pain (of body)	maumivu	usungu	14		v.
palm (of hand)	kiganja	chiganja / diganja	7 / 5		n.
person	mtu	munhu	1		n.

pestle	*mchi*	*mtwango*	3		n.
pot	*chungu*	*nyungu*	9		n.
quarrel	*ugomvi*	*ng'hondo*	9		n.
rib	*ubavu*	*mbafu*	?		n.
right hand	*mkono wa kulia*	*diganja dya kulume*	5		n.
rope	*kamba*	*sigi*	?		n.
saliva	*mate*	*mato*		6	n.
shoulder	*bega*	*yega, mayega*	5	6	n.
skin	*ngozi*	*ng'hingo*	9		n.
sweat	*jasho*	*fuke*	?		n.
tear(s)	*machozi*	*masosi*		6	n.
throat	*koo*	*kolomelo*	?		n.
to arrive	*kufika*	*kufika*	15		v.
to be married	*kuolewa*	*kutolwa*	15		v.
to beat	*kupiga kofi*	*kasugula kofi*	15		v.
to boil (intr.)	*kuchemka*	*kuchemsa*	15		v.
to breathe	*kupumua*	*kukeha*	15		v.
to burn up	*kuchoma*	*kulakasa*	15		v.
to carry water	*kuchukua maji*	*kusola meji*	15		v.
to chase away	*kufukuza*	*kuinga*	15		v.
to come	*kuja*	*kwija*	15		v.
to come from	*kutoka*	*kuleuka*	15		v.
to cook	*kupika*	*kutiga / kuteleka*	15		v.
to cool down	*kupoa*	*kuhola*	15		v.
to cover (a pot)	*kufunika*	*kugubika*	15		v.
to cut	*kukata*	*kukanha / kutema*	15		v.
to draw water	*kuteka maji*	*kunega meji*	15		v.
to extinguish	*kuzima*	*kusima*	15		v.
to fall	*kuanguka*	*kugwa*	15		v.
to fight	*kupigana*	*kwitoa*	15		v.
to fill	*kujaza*	*kumemesa*	15		v.
to finish	*kumaliza*	*kumala*	15		v.
to follow	*kufuatilia*	*kwandamila*	15		v.
to fry	*kukaanga*	*kukalanga*	15		v.
to go	*kuenda*	*kuluta*	15		v.
to go away	*kuondoka*	*kuleuka*	15		v.
to grind (grain)	*kusaga*	*kusaga*	15		v.
to hang up	*kutundika*	*kutundika*	15		v.
to hide (something)	*kuficha*	*kufisa*	15		v.
to insult	*kutukana*	*kuliga*	15		v.
to jump	*kuruka*	*kujumha*	15		v.
to kindle (fire)	*washa*	*kwasa (moto)*	15		v.

to lay down	kuweka	kuhika	15		v.
to leave	kuacha	kuleka	15		v.
to marry	kuoa	kutola	15		v.
to mix	kuchanganya	kugejageja	15		v.
to mould	kufinyanga	kuumba	15		v.
to pass	kupita	kuloka / kugenda	15		v.
to pound	kutwanga	kutwanga	15		v.
to pour	kumwaga / kumimina	kutila	15		v.
to put into	kutia	kuguma	15		v.
to quarrel	kugombana	kugoma	15		v.
to rest	kupumzika	kubwihila / kukwesela	15		v.
to return	kurudi	kuuya	15		v.
to run away	kukimbia	kubilima	15		v.
to seize	kushika	kukola	15		v.
to sew	kushona	kushona	15		v.
to shake (tr.)	kutikisa	kusingisa	15		v.
to sit down	kukaa	kukala	15		v.
to smash	kuponda	kuhonda	15		v.
to split (firewood)	kuchanja (kuni)	kutula ngodi / kuchanja	15		v.
to squat	kuchuchumaa	kusunyala	15		v.
to stand	kusimama	kuima	15		v.
to steal	kuiba	kuhija	15		v.
to sweep	kufagia	kushagila	15		v.
to take	kuchukua	kulonda / kusola	15		v.
to tear	kupasua	katula	15		v.
to throw	kutupa	kutaga	15		v.
to uncover	kufunua	kugubula	15		v.
to urinate	kukojoa	kukweja	15		v.
to wait	kungoja	kulindila	15		v.
to walk	kutembea	kugenda	15		v.
to wash clothes	kufua nguo	kufua suke	15		v.
tongue	ulimi	lulimi	11		n.
tooth	jino	gego, magego	5		n.
twins	mapacha	mhasa	10		n.
urine	mkojo	makwejo		6	n.
water	maji	meji	?		n.
wife	mke	mfele	1		n.
woman	mwanamke	mfele	1		n.
your father	baba yako	tate wako	?		n. + pron.

Appendix B: Identification sheets

1. Identification sheet Juma

Language: *Sagala [G39]*

a) Date of interview: *26-11-2016*

b) Place of interview: *Kilosa Mjinii, district: Kilosa District, country: Tanzania*

c) Consultant's name: *Juma Miraji Muhombolage*

d) Consultant's age (+ date of birth): *74, 1942*

e) Consultant's gender: *Male*

f) Consultant's birthplace: *village: Mkadage, district: Kilosa*

g) Consultant's geo-temporal track (places they have lived): *Kilosa, Dar es Salaam, visited many place during employment in Tanzania People's Defence Forces*

h) Consultant's first language: *Sagala*

i) Other languages spoken by consultant: *Swahili, some Sukuma*

j) Languages of both consultant's parents: *Sagala*

k) Languages spoken at consultant's hearth: *Sagala, Swahili*

2. Identification sheet Amina

Language: *Sagala [G39]*

a) Date of interview: *26-11-2016*

b) Place of interview: *Kilosa Mjini, district: Kilosa District, country: Tanzania*

c) Consultant's name: *Amina Abdallah Muhinakisambula*

d) Consultant's age (+ date of birth): *60, 1957*

e) Consultant's gender: *Female*

f) Consultant's birthplace: *village: Munisagala, district: Kilosa*

g) Consultant's geo-temporal track (places they have lived): *Munisagala, Dodoma, Mwanza, Iringa, Morogoro, Moshi, Kilosa Mjini*

h) Consultant's first language: *Sagala*

i) Other languages spoken by consultant: *Swahili, some Gogo*

j) Languages of both consultant's parents: *Sagala*

k) Languages spoken at consultant's hearth: *Sagala and Swahili*

3. Identification sheet Ndoweka

Language: *Sagala [G39]*

a) Date of interview: *19-11-16*

b) Place of interview: *Kilosa Mjini, district: Kilosa District, country: Tanzania*

c) Consultant's name: *Nasoro Muhamed Seleman Ndoweka Mkasi Chimalamwanza*

d) Consultant's age: *86*

e) Consultant's gender: *Male*

f) Consultant's birthplace: *Village: Tame, district: Kilosa*

g) Consultant's geo-temporal track (places they have lived): *Several places in Kilosa district: Tame, Muhangati, Msowelo, Rudewa Peapea, Kilosa Mjini, Mwasa, Kidodi, Kilombelo Sugar company, Mamboya, Idibo, Mikumi, Ulaya; has been to Dar es Salaam, Morogoro, Kilimanjaro region, Mahenge; has studied and worked in Kenya*

h) Consultant's first language: *Sagala*

i) Other languages spoken by consultant: *Sagala, Kaguru, Vidunda, Luguru, Kwere, Swahili, some English*

j) Languages of both consultant's parents: *Mother: Sagala, Kaguru; Father: Sagala (variety Zungwa), Luguru*

k) Languages spoken at consultant's hearth: *Sagala (variety Zungwa), Swahili*