AN INVESTIGATION INTO THE INTER-RELATEDNESS OF ILA, KAONDE, LENJE, NYANJA, SALA, SOLI AND TONGA

By

Avinat Mbwela Chitebeta

A Thesis Submitted to the University of Zambia
in Fulfillment of the Requirements of the Degree of
Doctor of Philosophy in Linguistic Science

THE UNIVERSITY OF ZAMBIA

LUSAKA

2019

COPYRIGHT

DECLARATION

I, AVINAT MBWELA CHITEBETA, do hereby declare that this thesis:

(a)   Represents my own work;

(b)   Has not previously been submitted for a degree at this or any other

University; and

(c)   Does not incorporate any published work or material from another thesis.

Signed:…………………………………………………………

Date:………………………………………………………….

APPROVAL

This thesis of **AVINAT MBWELA CHITEBETA** is approved as fulfilling the requirements for the award of the degree of Doctor of Philosophy in Linguistics by the University of Zambia.

Examiner 1

Name: _____   Signature_____   Date_____

Examiner 2

Name: _____   Signature_____   Date_____

External Examiner

Name: _____   Signature_____   Date_____

Chairperson - Board of Examiners

Name: _____   Signature_____   Date_____

Supervisor

Name: _____   Signature_____   Date_____

DEDICATION

To my husband, Jonas and children; Precious, Mutinta and Ethel.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABBREVIATIONS AND SYMBOLS

BB   Batu Botatwe

CDC   Curriculum Development Center

DEBS   District Education Board Secretary

IL   Ila

KA   Kaonde

LE   Lenje

MESVTEE   Ministry of Education, Science, Vocational Training and Early Education

NLF   National Literacy Framework

NY   Nyanja

PB   Proto Bantu

PT   Plateau Tonga

SA   Sala

SO   Soli

STLP   Strategy Towards a Language Policy

TG   Tonga

VT   Valley Tonga

ZEPH   Zambia Education Publishing House

ZL   Zambian Languages

DEFINITIONS OF TERMS

Selected terms drawn from the study are explained as follows:

(a)   **Proto-Bantu**: is a hypothesized mother of all Bantu languages.

(b)   **Bantu-Botatwe**: Literary meaning "three people," refers to a group of languages namely; Ila, Lenje, Lundwe, Sala, Soli, Tonga and Twa (Doke 1954).

(c)   **Comparative Linguistics**: A branch of linguistics which tries to find out whether languages are related by genetic descent or some other reasons.

(d)   **Historical Linguistics**: A branch of linguistics concerned with the reconstruction of the pre-historical languages with an aim of determining their relatedness in order to group them into language families.

(e)   **Inter-relatedness:** Relationships between languages.

(f)   **Lexicostatistics**: The statistical study of the basic vocabulary of two or more languages to determine their genetic relationship (Refer to Miti 1996:83).

(g)   **Swadesh word list**: A basic vocabulary word list of between 100 – 1000 words utilized in the lexicostatistical method utilized.

ABSTRACT

This study, **"An Investigation into the Inter-relatedness among Ila, Kaonde, Lenje, Nyanja, Sala, Soli, and Tonga"** involved a lexicostatistical survey on the inter-relatedness of the languages mentioned above. According to Doke (1954), Bantu Botatwe refers to the language group comprising the seven languages namely, Ila, Lenje, Lundwe, Sala, Soli, Tonga and Twa. While the affinity of these languages is generally accepted, no systematic studies of the lexicostatistical analysis have been made. The purpose of this study was to conduct a lexicostatistical survey on the degree of inter-relatedness of the languages investigated.

The study included Kaonde and Nyanja due to their geographical proximity to two Bantu Botatwe languages, Tonga and Soli, respectively. The study focused on Sala, one of the least documented languages in the Bantu Botatwe Group, to establish its degree of relatedness with Tonga, the language used for literacy and as a school subject in Sala-speaking areas. The study also focused on Soli, to ascertain the degree of inter-relatedness between Soli and Nyanja.

'**The Swadesh List**' of two hundred (200) vocabulary items, was the data collection instrument used. Three informants (translators) of above forty-five (45) years of age were carefully chosen from each language using purposeful sampling.

The data collected were analyzed using the lexicostatistical method whose results were presented in tables and graphs. Established cognates were coded: a plus (+) sign for cognates; a minus (-) sign for non-cognates; a hash (#) sign for the second set of cognates; a caret (^) sign for the third set of cognates where in a list of seven items three items were found to be cognates. Cognates were computed using a calculator to establish the percentage of inter-relatedness between the languages. The study established that Sala's degree of inter-relatedness with the other seven languages studied ranges from 39.5 to 81.5 percent. The highest percentage is between Sala and Lenje at 81.5 percent. The lowest is between Sala and Nyanja at 39.5 percent.

The study has also established that Soli and Nyanja are inter-related by 45 percent while Tonga and Kaonde are inter-related by 33 percent. The researcher recommends that: (1) A lexicostatistical survey be carried out in other Zambian languages. (2) A study be carried out to investigate the implications of the low percentages of inter-relatedness between Kaonde and Tonga in relation to the local language learners' performance. (3) Curriculum Development Center needs to develop materials for languages used for initial literacy e.g. Kikaonde in Mumbwa and Soli in Chongwe.

The researcher anticipates that the study will contribute to the existing comparative linguistic studies thereby enhancing knowledge and understanding by college or university lecturers as well as students.

Key words: Bantu Botatwe, Bantu Languages, Classification, Cognates, Comparative Linguistics, Inter-relatedness, Lexicostatistics, Percentages, Proto Bantu and Swadesh List.

CHAPTER ONE

INTRODUCTION

1.1   Background

This study attempts to ascertain the degree of inter-relatedness among some Bantu languages, namely; Ila, Kaonde, Lenje, Nyanja, Sala, Soli, and Tonga. The study involves a lexicostatistical analysis to establish the degree of inter-relatedness between the languages under study. The main focus of the study is on Sala, one of the least documented languages in the Bantu Botatwe Group, with the aim of trying to establish its degree of inter-relatedness with Tonga, the language used for initial literacy and as a school subject in the Sala speaking area. Further, the study aimed at trying to determine how closely Sala is related to the contiguous languages namely, Ila, Kaonde, Lenje. The study also focused on Soli, to ascertain its degree of inter-relatedness with Nyanja.

The study of the syntactic structures and morphosyntax of the Lenje and Tonga noun phrase [which the researcher did at the MA level] motivated the researcher to carry out further investigations on local languages and in particular in the area of lexical cognates. Further, the work of Hilary (2009) on lexicostatistical survey that included Tonga compared to non-Zambian languages motivated the researcher to investigate the degree of inter-relatedness among the seven Zambian languages dealt with in this study.

Torrend (1931: 83) states that Ila, Lenje, Soli and Tonga are part of the Bantu Botatwe group. Torrend also includes Lundwe and Twa in the Bantu Botatwe group. Fortune (1959:38) added Sala to the Bantu Botatwe group. However he does not indicate when Sala was added to it.

Miti (2006:47) states that Guthrie (1948) used the empirical method to differentiate Bantu languages (including the languages under study) which he termed *linguistic differentia,* set out as follows:

  i.   Lexical differentia – differences in vocabulary.

  ii. Grammatical differentia- differences in form and sentence structure.

  iii. Phonological differentia- differences of distinction between sound units.

  iv. Phonetic differentia – differences in actual speech sounds.

  v.   Tonal differentia – differences in tone systems.

According to Guthrie's (1948) classification of Bantu languages, Ila, Lenje, Sala, Soli and Tonga belong to the same group (Zone M). Kaonde is classified in zone L, while Nyanja is in zone N.

1.2   Brief Discussion of the Languages Investigated

The following discussion gives background information on each of the languages that have been investigated in this study. These include Ila, Kaonde, Lenje, Nyanja, Sala, Soli, and Tonga. The population figures used in this study were taken from the 2010 Census Report.

The languages dealt with in this study are all Bantu languages. The common characteristic of Bantu languages is that they use words such as *mu-ntu* or *mu-tu* for a human being or a person and the plural prefix *ba* refer to *ba-ntu* or *ba-tu*.

In Bantu languages every noun belongs to a class. The class is represented by a prefix as shown in the above paragraph. The prefix *mu-* represents class 1 in both Tonga and Ila, the plural prefix *ba-* indicates the change of class to class 2 although some nouns have a zero prefix.

Most Bantu words are made up of different syllables,every syllable ends with a vowel for example, languages dealt with in this study such as Nyanja, the noun 'eye' has the following two syllables: *di-so* – cv-cv; in Kaonde *ji-nso* meaning 'eye' is made up of cv-ccv. Sometimes Bantu words have a 'v' syllable in the onset especially in augument languages. For example, in Bemba *i-li-nso* meaning 'eye' starts with a vowel presented as v-cv-ccv. However, the languages studied are not augument languages. Some Bantu languages have consonant clusters than others as seen in the following words:

Tonga: (source: Ministry of Education (1977:4)

       Mbula 'wild fruit'   -   ccv-cv

        Nkaya 'bangle'   -   ccv-cv

        Nzala 'hunger'   -   ccv-cv

       Zyintu 'things'   -   ccv-ccv

       Bwizu 'grass'   -   ccv-cv

       Dwantuka 'to hop'   -   ccv-ccv-cv

Nyanja: (Source: Ministry of Education (1977:20))

       Kondwera 'rejoice' -   cv-cccv-cv

       Mphiripiri 'pepper' -   cccv-cv-cv-cv

       Mpholopolo 'ammunition'-   cccv-cv-cv-cv

       Kabvulumvulu 'whirlwind'- cv-ccv-cv-ccv-cv

Mnyamata 'boy'   -   cccv-cv-cv

Kaonde: (source: Ministry of Education (1977:60))

Kuzhika 'depth' -   cv-ccv-cv

Mwana 'baby'   -   ccv-cv

Mambo 'case'   -   cv-ccv

Bwela 'come back'-   ccv-cv

Michi 'medicine'-   cv-ccv

Among the examples given from the three languages (Tonga, Nyanja and Kaonde) Nyanja and Tonga have more consonant clusters than Kaonde.

Another common characteristic in Bantu languages is the aspect of reduplication. Reduplication refers to the repitition of a stem or of the whole word including the prefix. For example,

Ila: (Source: Interviews)   Tonga: (Source: intuition)

mbayimbayi 'soon',   linolino 'soon'

bucebuce 'bit by bit'   asyontosyonto 'bit by bit'

kumwetamweta 'smiling'   kusekaseka 'sociable'

mesomeso 'morally unstable'   mesomeso 'morally unstable'

1.2.1 Ila

Ila is a language spoken in the North-western part of the Southern Province, particularly the Bweengwa area in the Monze District, Namwala District and Mumbwa District of the Central Province. It is one of the Bantu Botatwe languages. According to Guthrie (1948), it belongs to zone M group 60 cluster 63. The Ila people share boundaries with the Sala people. The Census on Population and Housing report of 2010 shows that out of the total Zambian population of 13,092,666, 0.7 percent of the population use Ila as the predominant language of communication. This translates into 91,649 speakers. Ila is spoken in three provinces of Zambia with the following percentages: Central Province 2.7 percent out of 1,098,142 people in the province; Lusaka Province 0.1 percent out of 1,926,022 people in the province; Southern Province 3.7 percent out of 1,338,649 people in the province (See Appendix F).

1.2.2 Kaonde

Guthrie (1948) classified the Kaonde language under the Niger-Congo language phylum, under the Bantu languages in Zone L40. He says that this group is an off shoot of the Luba people of Congo, and is thus sometimes referred to as the Kaonde-Luba. From Congo, Kaonde people settled in some parts of the North-Western and central parts of Zambia. In the North-Western Province, they occupied Solwezi, Kasempa and Mufumbwe Districts while in Central Province they settled along the borders of Central and North-western Provinces in Mumbwa District. In the Central Province, the language shares its linguistic boundaries with Ila and Lenje within Mumbwa District. The Zambian map in Appendix E shows that in between Kaonde and Ila, there is a mixed language called Kaonde-Ila.

There are few linguistic texts that have been written on Kaonde these include the first bilingual English-Kaonde Dictionary written by Boroghall (1995) using the Kasempa dialect which is considered to be standard according to Mambwe (2008). The other text is English-Kaonde vocabulary authored by Wright (1985).

There are also a few literature books that have been published in and on Kaonde such as '*Kyapusana*' written by Mutembakyalo in 1988. The author named his book after the main character that went through life challenges.

Kaonde is one of the three regional lingua francas used for radio broadcasting, literacy and as a school subject in North Western Province particularly in Solwezi. The other two languages used within the Province are Lunda and Luvale. However, these two languages namely Lunda and Luvale are not part of this study.

According to the Census on Population and Housing report of 2010, Kaonde is largely spoken in North-Western Province by 29.6 percent out of 596,860 people in the province, 0.7 percent out of 1,741,192 people in the Copperbelt Province, 0.9 percent out of 1,098,142 people in Central Province, 0.2 percent out of 1,926,022 people in Lusaka Province, 0.3percent out of 737,287 people in Western Province and 0.1 percent out of 1,338,649 people in Southern Province. However, out of the total population of Zambia (13,092,666) in the year 2010, there were 1.9 percent of people using Kaonde as their predominant language of communication. (see Appendix F). This translates into 248,761 speakers.

1.2.3   Lenje

According to Torrend (1931:83), Lenje is "one of the indigenous languages spoken by the people of Central Province of Zambia." According to the Census on Population and Housing Report of 2010 out of 1,098,142 people in Central Province, 10.4 percent speak Lenje; 0.1 percent on the Copperbelt Province speak Lenje out of 1,741,192, while 0.6 percent speak Lenje out of 1,926,022 people found in Lusaka Province, and 0.1 percent of Lenje speakers are found in the Southern Province where there is a total population of 1,338,649 people.   Out of the total Zambian population of 13,092,666 in the 2010 Census, only 1.2 percent spoke Lenje as their predominant language of communication by residence (see Appendix F). The percentage translates into 157,112 speakers.

Lenje belongs to Zone M group 60 cluster 61 according to Guthrie's (1948) classification. The language is related to other languages within the Bantu Botatwe group in which Doke (1954) divided into seven diferrent languages namely, Ila, Lenje, Lundwe, Sala, Soli, Tonga, and Twa, as in Figure 1 below although this study does not include Lundwe and Twa due to the narrowing of the topic.

Bantu Botatwe

Ila   Lenje   Lundwe   Sala   Soli   Tonga   Twa
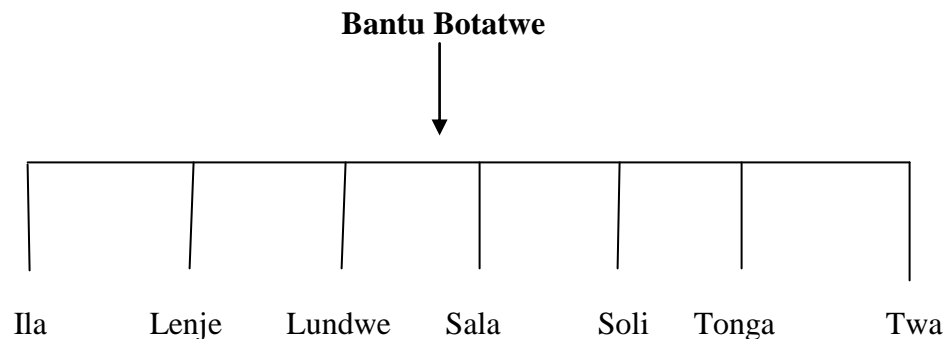
Figure 1: Bantu Botatwe Languages (Source: Doke 1954)

The oral interviews carried out by the researcher in March 2011 in the city of Lusaka show that generally, some non-Lenje speaking people believe that Lenje is a dialect of Tonga. However, the classification done by Guthrie (1948) suggests that Lenje and Tonga are related but different. The difference between these two languages can be noted using Guthrie's linguistic differentia, particularly phonological differentia which looks at differences of distinction between sound units. For example, *shoonse* 'all' in Lenje and *zyoonse* in Tonga; *kucata* 'catch' in Lenje and *kujata* in Tonga. Instead, Guthrie (1967:81) identifies Toka and Leya as dialects of Tonga with the following classifications: Tonga M64a, Toka M64b and Leya M64c.

1.2.4   Nyanja

According to Guthrie's (1948) classification of Bantu languages, Nyanja falls under Zone N, which is made up of four groups, namely 10, 20, 30 and 40. Nyanja belongs to Zone N, Group 30 as a dialect cluster called N31. According to Guthrie (1948), Zone N comprises three dialects, namely Nyanja, Cewa, and Manganja identified as N31a, and N31b, and N31c respectively. Ohannessian and Kashoki (1978:401) show that "Nyanja is a regional lingua franca used in Eastern and Lusaka Provinces." According to Miti (1988), Chewa, Chinsenga and Chingoni are the three varieties referred to collectively as Zambian 'Chinyanja.' However, Ngalande (2007) argues that there is no language called Nyanja or a speech community referred to as Nyanja. What is known as Nyanja, referred to and spelt as 'Cinyanja,' is Chewa, a native language of the Chewa people of Katete District whose Paramount Chief is Gawa Undi. Ngalande (2007) further notes that a section of speakers in Lusaka speak some form of a lingua franca that has a reasonable vocabulary from Chewa and is called 'Cinyanja.' Nyanja is also spoken in other districts in the Eastern Province;

these include the districts neighboring Katete, namely, Mambwe and Chipata. Others are Petauke, Lundazi and Chadidza. Nyanja is also spoken in some of Zambia's neighboring countries, for example, Malawi where it is officially termed 'Chewa;' Mozambique, Tanzania and Zimbabwe. While Ngalande argues that there is no speech community referred to as Nyanja, it is a known fact that among the seven Zambian languages used for literacy in schools, Nyanja is one of them.

Zambia has radio and television broadcasts in Nyanja. There was also a monthly publication called *Tsopano* which is currently out of circulation, while Malawi has radio broadcasts in Chewa and a publication called '*Boma Latha'*, a Chichewa Daily.

According to the 2010 Census on Population and Hosing report, Nyanja is the second predominant language of communication by residents in Zambia. Out of the total population of 13,092,666 in the year 2010, at least 14.8 percent of the population in Zambia, which translates into 1,937,715, spoke Nyanja also the greatest number of second language speakers.

During the discussions between the researcher and Mr. Mundia Mwendende of the Department of Languages at the Curriculum Development Center in March 2011, it was confirmed that Nyanja is used for literacy and as a subject in schools in Lusaka Province, comprised of Chongwe District, Kafue District, Lusaka District and Luangwa District. This means that Soli children in Chongwe District are learning in Nyanja, which is not their mother tongue, for their literacy programs at school. This provides the motivation for including Nyanja in the current study in order to discover its degree of inter-relatedness with Soli. It is hoped that the results of this study will help to assess whether the

established degree of inter-relatedness between these two languages shows the relevance of the use of Nyanja for literacy in Chongwe District.

1.2.5   Sala

Sala is a language spoken along the boundaries of Central, Lusaka and Southern provinces, for example, Lusaka-west, Nampundwe and Shibuyunji. Fortune (1959:38) states that "Sala was added by Doke (1954) to the Bantu Botatwe group sometime after Guthrie's classification." The preliminary investigations carried out by the researcher with Mr. Mwiimbu, Headman of Kambya Mwiimbu Village, Chief Shakumbila, in the presence of a group of his subjects in March 2011, stated that

> "The Sala people originated in Gwembe. One man by the name of Choongo and three ladies (wives) namely, Chintu, Mungalu and Maninga moved from Gwembe due to inter-village fightings and settled along Kafue River. Later on these four people crossed the river and settled in Chipapa area. After some years, Chipapa area became small, causing this family to move to Lilayi where there was more space for farming. From Lilayi, the family grow bigger and space for crop farming and pasture for their animals became limited. Therefore, Choongo and his wives moved to Farmers area along Mumbwa Road, while others moved to Chabota area and from Chabota others moved on to Shabasonke" (Oral interviews: Underson Mwiimbu, Chief Shakumbila, Mumbwa District.)

On the other hand, one person who was interviewed within Chief Shakumbila's chiefdom believed and stated that "Salas were originally Lenjes. According to this belief, at one point one group decided to separate themselves from the Lenjes, and the Lenjes called the group 'basala' meaning 'they have chosen to separate.' Since then they continued to call them 'ba Sala'." (Oral interviews: J. B. Mwinga).

During the oral interviews by the researcher, Chief Shakumbila, stated that "the geographical area occupied by the Sala people stretches from Landless Corner on the Great

North Road through Namakolongo in the north. From there, it goes up to Kabile, via Mashili Dam, Mwembeshi River in the north east, and passes through Makombwe stream and Nsanje Hill in the east. Kapili Kamamvwa is in the south east while in the south the area encompasses Kafue plains, Chise, Kabulungwe and Masene. In the south the area covers Kakombwe Hill, Makombe Dam and crosses the old Mongu road and further joins Kabwe road at four miles in the north (see Appendix D).

The nearest church (Shakumbila Seventh-day Adventist Church) to the chief's palace investigated during the survey shows that Sala people use Tonga Bibles and Hymn books, while discussions are conducted in both Sala and Tonga. Tonga is used in schools. Sala is surrounded by other languages such as Ila, Lenje, Kaonde-Ila, Kaonde, and Tonga. We did not come across any records on how many speak Sala from the 2010 Census on Population and Housing report.

1.2.6 Soli

Soli is one of the indigenous languages used by people in Chongwe District and around the borders of Chongwe and Kafue District in Lusaka Province. The Soli people are surrounded by several language groups namely: Chikunda, spoken in the eastern part of the Soli chiefdom along the borders with Luangwa; Ila and Kaonde on the western border with Mumbwa; Lenje and Sala on the Northern boarder with Kabwe; Tokaleya on the southern side of Chirundu; and Tonga, on the southern side along the borders with Kafue.[1] Torrend (1931:83) states that, "Soli belongs to Zone M, group 60, cluster 62." In the Census on

---

[1] (Oral interviews: with three senior citizens, namely, Mr. Phinias Mukuwa, Mr. Silas Njoomwa and Mr. Mumba Lupekesa).

Population and Housing 2010 report there is no specific information given on the number of Soli speakers.

The authorities in the Department of Languages at the Curriculum Development Centre confirmed that the language used for literacy campaigns in the Soli speaking area is Nyanja. It is also used as a content subject in schools within Chongwe District (oral interviews: Mr. Mundia Mwendende).

Recently, after studying the literacy situation and establishing the factors that affect educational quality in Zambia, the Ministry of Education, Science, Vocational Training and Early Education (MESVTEE), through the Curriculum Development Center (CDC) developed a 'Strategy Towards a Language Policy (STLP)' as shown in Table 1.

Table 1: Proposed Language of Instruction in Zambia. (**Source**: National Literacy Framework (2013 page 5-6).

Grade	Content Subjects and Literacy	Language of Instruction
1	All learning areas	Local Languages
2	All learning areas	Local Languages
	Content subjects and Literacy in ZL	Local Languages
	English Language and Oral Literacy	English Language
3	Content subjects and Literacy in ZL	Local Language

	English Language and Literacy	English Language
4	Content subjects and Literacy in ZL	Local Language
	English Language and Literacy	English Language
5-7	Content subjects	English Language
	English	English Language
	Zambian Languages	Local Languages

In Table 1 above, the purpose of MESVTEE is to "introduce instruction in a familiar language so as to build learners' arsenal for learning to read in other languages as well as learning content subjects." (National Literacy Framework NLF 2013 page 5.) The proposed language of instruction in Zambia will help the learner from Grade one to Grade four to learn all learning areas, content subjects and literacy in Zambian languages (ZL) using local languages, except English language subject and Oral Literacy, which will use English language for Literacy. Then from grade five to seven, other content subjects and English language subject will all be taught in English language except Zambian language subjects will be taught in local languages.

Oral interviews carried out in February 2012 by the researcher through the former District Pastor for Chilanga Mission District, Pastor Jonas Chitebeta confirmed that at Chipapa Seventh-day Adventist Church, Soli is one of the main languages used for preaching and Bible studies. Investigations have also shown that there are many Soli speakers in the

Chipapa area within Kafue District. However, Nyanja is used mainly in the Bible readings but discussions and preaching are conducted in Soli with Nyanja interpretations. Nyanja is also used for literacy and as a school subject in Chisankane area in Kafue District where there are many Soli schoolgoing children. The researcher attended a church service at Chisankane Seventh-Day Adventist Church next to Chisankane Basic School where it was observed that the language used for preaching and Bible study is Soli with Nyanja interpretations. The researcher also observed that there is language mixing within Chisankane area due to language contact resulting from some of the Tonga and Sala families who have settled in the area. It is interesting to note that the Soli word 'Chisankane' is translated as 'mixing.' The researcher believes that the name 'Chisankane' fits the area because of the different languages used in the same area, namely: Nyanja, Sala, Soli and Tonga.

According to preliminary interviews carried out in March 2011 by the researcher at Chongwe market with a group of five people led by Mr. Phinias Mukuwa from Chibwalu village in Chongwe, Nyanja is the language of communication in Chongwe town while Soli is commonly used by the Solis amongst themselves in villages. Solis are found within Lusaka Province especially in Chongwe, Shikabeta, Mumpasha, Bundabunda and Kafue (oral group interviews). The recent oral interviews conducted in Chongwe District, on 20th February, 2013, between the researcher and Mr. John Takile, Senior Headman Kakuka, revealed that Sun Radio in Lusaka has Radio Programs in Soli.

Origin:

The origins of the Soli people according to the study carried out by Manchishi and Musona (undated:7-8) shows that "traditionally, the Soli say they came from Chief Mwata Yamvo (ku Buluba). They are part of the earliest migrations of the Nyangu (Nyendwa) clan… The study reveals that Soli people came in the company of the Lala but later on, broke away and left for the present Soliland."

According to Breslford (1956:63) "the Soli are a branch of the Luba migrations that came in from the west, a branch that includes the Kaonde, Lala, Lamba and other tribes that came into the territory along the route south of the Luapula and Bangweulu." Fagan (1966) also confirms the relationship in origin between the Soli and Lala, he says that the Soli also seem to have been influenced by a large group of Luba derived people to the north and north-east, the traditions of this group which includes the Lamba, Ushi and Lala, seem to link them in a common process of migration and settlement.

1.2.7 Tonga

Tonga belongs to the Bantu Botatwe Group, which is classified in the Central Zone according to Torrend (1931) and is the main language in that zone. Tonga is one of the indigenous languages spoken by the people of the Southern Province of Zambia. It belongs to Zone M, Group 60, cluster 64 according to Guthrie's (1948) classification. Guthrie (1967:81) adds that Tonga has the following three dialects: "Tonga (M64a) Toka (M64b), and Leya (M64c)."

Kashoki (1978) has shown that Tonga is generally used for literacy and taught as a subject in schools in Southern province apart from Livingstone where Lozi was used in some

schools as a content subject at the time of Ohannessian and Kashoki's research. However, current investigations carried out by the researcher in Livingstone District have shown that Lozi is no longer used as a content subject within the district. Following the new language policy of using a familiar language for instruction, Tonga is used for content subjects and literacy in ZL from grade one to grade four and as a content subject from grade 5-7 throughout Livingstone District.

According to the Central Statistics report of 2010, Tonga is the third predominant language of communication by residents in Zambia. Out of the 13,092,666 total population of Zambia in 2010, 11.4 percent spoke Tonga. This translates into 1,492,564 speakers.

The oral interviews held by the researcher on 23rd May 2012 at the Ministry of Education, Provincial Headquarters in Livingstone, Mrs. R. M. Chulu, the Senior Educational Standards Officer, confirmed that currently Tonga is used for literacy and taught as a subject in all schools in Southern Province. However, there are a few schools in Kazungula District where Lozi is used as a familiar language, as alluded to earlier in the new language policy that learners should use local languages which are familiar for content subjects and literacy in ZL. For example, Riverview Basic School and Mambova Basic School are some of the few schools using Lozi as the medium of instruction according to the new language policy. Further, the Curriculum Development Center confirmed that Tonga is one of the Zambian languages used for broadcasting for both radio and TV programs. It is also used for literacy campaigns and as a content subject in Choma, Kalomo, Kazungula, Livingstone, Mazabuka and Monze districts as well as Kabwe rural in Central Province. Kabwe urban and some districts within Central Province such as Kapiri Mposhi and Serenje use Bemba for literacy campaigns and as a content subject.

According to Kashoki (2002), Tonga covers a number of varieties spoken mainly in Southern Zambia by people who call themselves *ba-Tonga* while Sibajene (2013:1) added that Tonga is also spoken "in some parts of Zimbabwe especially around the Gokwe North, Gokwe South, Kanyi, Nyaminyami and Victoria Falls area." Within the Tonga area (*bu-Tonga*) there are dialectal differences in the sound system and vocabulary. The greatest differences are between the north comprising of Mazabuka, Monze, Choma and Kalomo districts and south comprising of Siavonga, Gwembe and Sinazongwe districts where the so-called 'Plateau Tonga (PT)' and Valley Tonga (VT)' varieties respectively are found, although there is some east-west variation as well. Sibajene (2013:1) points out that the dialectal variations are "at phonological, morphological, syntactic and lexical levels." These different dialects often use the same spelling to represent different sounds, for example, the spelling *sy* represents southern [sy] and northern [hy], and the latter sometimes spelt *h*. Both Carter (2002) and Hopgood (1992) admit that there are pronounciation variations between Plateau Tonga and Valley Tonga. This study used the approved Tonga orthography to avoid the challenges of the dialectal differences found in the Tonga sound system.

Generally, some people believe that Lenje is a dialect of Tonga. However, Guthrie's (1948) classification suggests that Lenje and Tonga are related but distinct languages.

Of the languages being investigated, three are some of the languages that have been chosen for use as lingua francas alongside English in the media, for public education, school subjects and languages of literacy in schools. These are Kaonde, Nyanja and Tonga. The other languages that have been chosen by the Zambian government for this purpose are Bemba, Lozi, Lunda and Luvale.

1.3   Statement of the Problem

Generally, "*Bantu Botatwe*" literally means "Three People" believed to comprise the Tonga, Lenje and Ila. However, Doke (1954) indicates that there are more languages in the Bantu Botatwe Group and has divided them into seven languages namely: Ila, Lenje, Lundwe, Sala, Soli, Tonga and Twa. Originally three languages comprised the Bantu Botatwe group but the concept has continued even when other linguistic groupings have been included in the group.

However, of late there have been some Lenje speakers advocating that Lenje children should be taught Lenje in schools and not Tonga. Their arguments have been that Lenje is a distinct language and that Lenje is not a dialect of Tonga (oral interviews with L. Mwesa, Administrative Assistant- Rusangu University, Monze and P. C. Kachenga, Retired Teacher- Libala, Lusaka).

Amongst the so called Bantu Botatwe Languages (BBL) Sala and Ila are among the least documented. So far, the investigations carried out by the researcher in March 2011 at Shakumbila, Mwembeshi, Lusaka, the Central Statistics Office, and Curriculum Development Center have shown that there is no comprehensive study that has been carried out to determine their status in this group and the levels of their inter-relatedness with the other languages in the same group. It is the intention of this study to fill this gap.

It has been mentioned in the introduction that Tonga is used for radio broadcasting and television programs in Zambia. It is also used for literacy campaigns and as a school subject in schools within the Sala speaking area of Mumbwa District. The challenges faced by Sala learners are not known. The determination of the relatedness between Sala and

Tonga will help to establish whether there is justification in the use of Tonga as a school subject in the Sala speaking area.

As has already been mentioned in the introduction, Nyanja is used for literacy and as a school subject in the Soli speaking areas of Chisankane and Chongwe. The determination of the inter-relatedness of these languages may provide insights into some of the challenges that could be encountered by Soli learners taking Nyanja as a school subject.

In view of these controversies with regard to the Bantu Botatwe Group and the unclear relationship between this group and other languages associated with them, such as Kaonde, Nyanja and Sala, an investigation on the degree of their relatedness provides very useful information. Also, preliminary investigations have shown that so far there has been no study carried out to determine the degree of inter-relatedness between, Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga. Therefore, this study provides a valuable contribution to the field of comparative Bantu linguistics.

Guthrie (1948) classified Ila, Lenje, Sala, Soli and Tonga as belonging to Group M, as mentioned in the introduction under section 1.0 page 2. He used the empirical method such as differences in vocabulary, differences in form and sentence structure, differences in sound units, differences in actual speech sounds and differences in tone systems (Miti 2006). However, He did not establish the percentage of the degree of inter-relatedness between these languages and the regional languages used in schools for literacy and as school subjects such as Nyanja which is used in the Soli speaking area and Tonga which is used in Kaonde speaking areas. This study has been carried out in order to bridge this gap in the field of comparative Bantu linguistics.

1.4   Purpose of the Study

The purpose of this study is to determine the degree of inter-relatedness among Ila, Lenje, Kaonde, Nyanja, Sala, Soli and Tonga. The specific objectives of the study were;

(a) To establish the degree of inter-relatedness between Sala and the other Bantu Botatwe Languages namely, Ila, Lenje, Soli and Tonga in Mweembezhi;

(b) To ascertain how closely related Soli is to Nyanja;

(c) To ascertain how closely related Kaonde is to Tonga;

(d) To ascertain whether there is any retention of the Proto-Bantu vocabulary in the languages under investigation.

1.5   Research Questions

The study attempts to answer the following questions:

(a) What is the degree of inter-relatedness between Sala and the other Bantu Botatwe languages namely: Ila, Lenje, Soli and Tonga.

(b) What is the degree of inter-relatedness between Soli and Nyanja?

(c) What is the degree of inter-relatedness between Kaonde and Tonga?

(d)   Is there any retention of the Proto-Bantu vocabulary in the languages under investigation?

1.6   Significance of the Study

The importance of the study is summarized in the following underlying reasons:

1.5.1 It is intended to bridge the gap that exists in linguistic studies especially in ascertaining the degree of inter-relatedness of the languages investigated namely: Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga.

1.5.2 The results may be used by linguistic researchers as a stepping stone for further studies in Bantu comparative linguistics.

1.5.3. It will be made available to language policy makers to offer advice as they evaluate Zambian language policies.

1.5.4. It will make a contribution to existing knowledge in comparative linguistics for Bantu languages of Zambia.

1.7   Analytical Framework

This study employed the tenets from historical, comparative linguistics and lexicostatistics. Historical linguistics compares living languages whereas comparative linguistics utilizes the field of historical linguistics. Matthews (2005) defines historical linguistics as the study of change in individual languages and in language generally. He further states that historical linguistics is distinguished by most schools of structural linguistics as a branch of the subject concerned with diachronic relations among language systems, separate from the findings of synchronic or descriptive linguistics. Grimm (1822), the historical linguists noticed recurrent correspondences between the sounds of cognate words in the early Indo-European (IE) languages. They explained these historical sound changes or "sound laws." One of the first sound laws to be discovered was the Germanic consonant shift "Grimm's Law", which converted earlier voiceless stops to voiceless fricatives (cf. Sanskrit *trayas*: English **th**ree), voiced stops to voiceless stops (Sanskrit *dvau* : English. *two*),and "voiced aspirates" to plain voiced stops in Germanic (Sanskrit **bh**ra tar- : English **b**rother). As more and more sound changes were studied, an important generalization emerged: if the stable, language-specific phonetic environment for a given sound change was satisfied, the

21

change took place. It was observed that the change of voiceless stops to voiceless fricatives in Germanic, for example, always applied word-initially and after vowels and consonants, but never after stops or fricatives (Skt. **st**ar- : Eng. **st**rew, not **sth**rew).The global claim that "sound change is regular," was first made by the German "Neogrammarian" (Junggrammatiker) school in the late 1870's and has been accepted in some form ever since.

According toJaggar (2010), the fundamental technique of comparative linguistics is to compare phonological systems, morphological systems, syntax and the lexicon of two or more languages using techniques, such as the comparative method. During the 19th century, the theory of comparative method was developed through the study of the Indo-European languages, and it remains the typical technique by which conventional linguistics evaluates whether two or more languages are related.

Campbell (2004:111) affirms that

> "the comparative method is a technique used by linguists to demonstrate genetic relationships between languages. It aims to confirm whether two or more historically attested languages are descended from a single proto-language by comparing lists of similar terms. From these similar lists, regular sound correspondences between the languages are recognized, and a series of regular sound changes can then be postulated which allows the proto-language to be reconstructed from its daughter languages. A correlation is deemed certain only if an incomplete reconstruction of the antecedent is practical and if regular sound correspondence can be recognized with chance similarities ruled out."

Miti (2006) discusses the comparative method and states that it must take into account the following procedure that involves four major steps:

(a) Collection and organization of data from the languages which are suspected to be genetically related. The data should consist of cognate[i] sets.

(b) Elimination of suspicious-looking items such as borrowings and chance resemblances. For example, the ciCewa word *dothi,* which means 'soil/earth' in English. Some people have suggested that the ciCewa word is a borrowing from the English word *dirt.* The word *dirt* was not imported in the ciCewa area, so it is unlikely that ciCewa would borrow the word. From the phonological point of view, there is no reason the vowel in the word dirt should be rendered as /o/ in ciCewa and not /e/. This is an example of chance resemblances that need to be eliminated from any comparative data (Miti 2006).

(c) Determining of sound resemblances that exist between sounds in the same position of each set of cognate items in the language under comparison.

Mkude (1995:7) states that "Comparative Linguistics recognizes some strikingly systematic similarities and variations among human languages." A number of theories have been proposed to account for such similarities and variations. He further noted that sometimes there are remarkable similar features between languages which are not genetically or geographically related where the languages concerned would have historically influenced one another. A good example from Mkude's study in relation to English and Swahili is in the following pairs of words (p.17):

Swahili	English
shati	shirt
bunda	bundle
gauni	gown (p.17)

The examples shown above reveal how the Swahili language has added to its native stock a substantial amount of vocabulary drawn from English sources.

Hinnebusch (1999:174) states that "the comparative method is the procedure by which form-meaning elements of related or potentially relatable languages are matched in sets of correspondences in order to establish their genetic relatedness." The comparative method remains the standard method by which conventional linguists judge whether two or more languages are inter-related. The current study seeks to establish the degree of inter-relatedness among eight Zambian languages (Ila, Kaonde, Lenje, Nyanja, Sala, Soli, and Tonga) by applying the lexicostatistical technique whose base is comparative linguistics. The lexicostatistical techniques involve quantitative comparison of lexical cognates. The study utilizes the basic vocabulary list called the Swadesh word list.

The Swadesh word list is a list of basic vocabulary of about one hundred to one thousand words used in comparative linguistics. Swadesh lists are usually used in lexicostatistics (quantitative language relatedness assessment) and glottochronology (language divergence dating). The Swadesh word list provides a better yardstick in identifying lexical variations because it contains all basic words that are used to express universal concepts. It is chosen for its universal, cultural independent, availability in as many languages as possible.

1.7.1   Lexicostatistics

Lexicostatics is an approach to comparative linguistics that involves the quantitative comparison of lexical cognates. Lexicostatics is related to the comparative method but does not reconstruct a proto-language. It is to be distinguished from glottochronology, which attempts to use lexicostatical methods to estimate the length of time since two or more

languages diverged from a common earlier proto-language. This is merely one application of lexicostatics, however, and other applications of it may not share the assumption of a constant rate of change for basic lexical items.

The term "lexicostatics" is misleading in that mathematical equations are used but not statistics. Other features of the language may be used other than the lexicon, though this is not usual. Whereas the comparative method used shared identified innovations to determine sub-groups, lexicostatistics does not identify these. The latter is the distance based method but the comparative method considers language characters directly. The Lexicostatics method is a simple and fast technique relative to the comparative method but has limitations. It can be validated by cross-checking the trees produced by both methods.

Historically, Lexicostatistics was developed by Morris Swadesh in a series of articles in the 1950s, although the concept seems to have been proposed in 1834 by Dumont d'Urville, who compared various Oceanic languages and proposed a method for calculating a coefficient of relationship. Hymes (1960) and Embleton (1986) review the history of lexicostatistics.

Isidore Dyen is a leading exponent of the lexicostatistics application. He used lexicostatistics to classify Austronesian languages as well as Indo-European ones. A major study of the latter was reported by Dyen, Kruskal and Black (1992). Studies have been carried out of Amerindian and African Languages.

Gudschinsky (1956) studied the ABC's of lexicostatistics and observed that by simple inspection of comparable world lists, for example, the fact of the relationship of closely related languages can be discovered. However, on the basis of simple inspection no one

can precisely say how closely related two languages are. In Glottochronology, Gudschinsky (1956) came up with basic assumptions of lexicostatistics. The first assumption is that some parts of the vocabulary of any language are assumed on empirical evidence to be much less subject to change than other parts. These basic core vocabularies which include such items in the material culture, are frequently, and borrowed along with the cultural items. Such terms are also easily lost with a change in the material culture, or the borrowing of the new items. The contrast between the basic vocabulary and general vocabulary may be seen in the illustrations of French loan words in English.

> As against perhaps 50 percent of borrowed correspondences between English and French in the general vocabulary, we find only six percent in the basic vocabulary. Residual correspondences are found to be 27 percent. Thus, the archaic residuum after 5 000 years turns out to be five times greater than 2 000 years of accumulated borrowings (Swadesh [1951a] p.13).

The second basic assumption of lexicostatistics is that the rate of change of vocabulary items in the basic core of relatively stable vocabulary is constant through time. That is, given the basic number of words in a certain language a certain percentage of these words will remain unchanged after a second thousand years, the same percentage of words will remain unchanged at the end of the second period (Lees, 1953).   This span does not provide adequate evidence for a constant rate of loss over a long period of time, since the assumption has not yet been checked for a time span greater than 2,200 years.

This study is not concentrating on the rate of relatively stable vocabulary that are constant through time, but the study used lexicostatistics to identify lexical cognates in order to establish the degree of inter-relatedness of the languages under study. Bynon (1977:271) states that lexicostatistics may be useful "… as a preliminary to reconstruction only." This

means that lexicostatistics may be useful as a preliminary to the reconstruction of languages.

According to Batibo (1998a) the aim of lexicostatistics is to generate a list of universal culture-free forms. Words are then collected and equivalent meanings of each language are established. This method uses quantitative evaluation where collected list of vocabulary is quantified according to points. Then cognate words are counted and recorded in percentage form to show the degree of inter-relatedness between the languages under review. Swadesh reduced a larger set of meanings down to 200 originally. He later found that it was necessary to reduce it further but that he could include some meanings that were not in his original list, giving his later 100-item list. The Swadesh list in Wikitionary gives a total 207 meanings in a number of languages. Alternative lists for particular purposes have been generated, for example, Dyen, Krustal and Black (1992) have 200 meanings for 84 Indo-European languages in digital form.

Additional information on lexicostatistical theory has been given by Blust (1981), who makes a useful distinction between horizontal lexicostatistics and vertical lexicostatistics. He states that the horizontal lexicostatistics is a technique that compares lexical data from different languages which are supposed to have been attested in the same time period and to be roughly contemporaneous. The vertical lexicostatistics technique compares lexical data from an earlier stage of a particular language with data from other languages which are assumed to be descendants from this language.

1.7.2. Determining cognates

Determining cognates requires a trained and experienced linguist. However it should be noted that the decisions to determine cognates may need to be refined as the state of knowledge increases. On the other hand, lexicostatistics does not rely on all the decisions being correct. For each pair of lists, the cognancy of the form could be positive, negative or determinate. Sometimes a language has two words for one meaning, for example: *small* and *little* for *not big*. This study was guided by trained and experienced linguists, Dr. Nkolola M. Wakumelo (deceased) and Dr. J. Lubinda.

This study employs the lexicostatistical techniques to determine the inter-relatedness of eight Zambian languages. It uses a basic vocabulary of 200 words. These languages are Ila, Kaonde, Kaonde-Ila, Lenje, Nyanja, Sala, Soli, and Tonga. The study also uses Batibo's (1998a) and Miti's (1996) methods of collecting words and establishing meanings in each language. This method employs quantitative evaluation where collected rural form vocabulary is quantified according to points. Then cognates are counted and recorded in percentage form to show the degree of inter-relatedness. The study employs horizontal lexicostatistics.

1.7.3. Language Family Tree

Creation of the language tree is based on various sub-grouping methods. This study used Dyen, Krustal and Black's subgrouping method that places all lists in a pool. The two closests members are removed and form a nucleus which is placed in the pool. Then this step is repeated. Under certain conditions a nucleus becomes a group. This is repeated until the pool only contains one group. Then calculations need to be made of nucleus and group lexical percentages.

1.7.4. Criticism aganist Lexicostatisdtical method

Hijer (1956) criticised the lexicostatistical method and has shown that there were difficulties in finding equivalents to the meaning of items while many have found it necessary to modify the Swadesh's lists. Gudnchisky (1956) questioned whether it was possible to obtain a universal list. Sometimes lexicostatistics has been used with lexical similarity being used rather than cognance to find resemblances. The choice of meaning slots is subjective as is the choice of synonyms.

1.8 Limitations of the Study

The researcher could not travel to North-western and Eastern Provinces bacuase Kaonde is also spoken in Mumbwa while Nyanja is also spoken in Lusaka respectively.

1.9 Scope and Delimitations of the Study

This study is delimited to eight Zambian languages; one from Lusaka Province, namely, Soli; three from Central Province, namely, Lenje and Sala; one from North-western Province- namely Kaonde, which is also spoken in Mumbwa District of CentralProvince (See Appendix E); one from Eastern and Lusaka Provinces, namely, Nyanja/Cichewa; two from Southern Province, namely, Ila and Tonga.

According to Miti (2006), different methods have been used to classify languages, namely, alphabetical, areal, typological, political and genetic classifications. The current study is delimited to typological classification since its aim is to determine the degree of inter-relatedness of languages and not reconstruction of the languages. Song (2011) defines linguistic typology as a subfield of linguistics that studies and classifies languages according to their structural features. Its aim is to describe and explain common properties

and the structural diversity of the world's languages. It includes three sub-disciplines: qualitative typology, which deals with the issue of comparing languages and within language variance; quantitative typology, which deals with the distribution of structural patterns in the world's languages; and theoretical typology, which explains the distributions of structural patterns in the world's languages. This study has utilised the qualitative typology, where Ila, Kaonde, Lenje, Nyanja, Sala, Soli, and Tonga have been compared.

Different methods of determining relationships between languages have been cited by some scholars. For example, Batibo (1998b:16) states that there is "the method of resemblance, the comparative method, the lexicostatistical method and the method of structural comparison." This study employs only the lexicostatistical technique, which uses vocabulary cognates to determine the degree of inter-relatedness between languages. The study is based on the basic vocabulary of the languages under review.

This research covers eight Zambian languages which are widely distributed in Central province, Eastern province, Lusaka province, North-western province and Southern province. Although the researcher could not travel to all the locations as aluded to in 1.8 above, she overcame this challenge by interviewing and distributing the Swadesh list to native speakers of the specific languages that were available at Rusangu University (formally Zambia Adventist University) in Monze, University of Zambia, Lusaka; City market, Lusaka; Chibwalu village in Chongwe District, Situmbeko Village on Mumbwa road, Chief Mumba's Palace and surrounding villages, Mumbwa township, Senior Chief Shakumbila's Palace and the surrounding villages, and Rockview Institute on Mumbwa Road near the Mwembeshi satellite. The researcher made a deliberate choice to give

instructions, teaching how to use and distribute the Swadesh list in both urban and rural areas in order to check whether there were vocabulary differences between rural speakers and urban speakers that affected the lexical items of the languages under review.

1.10   The Structure of the Thesis

The thesis comprises five chapters. Chapter One includes the introduction, background information on the languages under study, structure of the thesis, statement of the problem, purpose of the study with specific objectives, research questions, significance of the study, theoretical and analytical framework, operational definition of terms, scope and limitations of the study. The second chapter hubs on literature review. Chapter Three describes the methodology, which includes the research design, study location, study population, sampling techniques, data gathering techniques and, ethical considerations, among others. The fourth chapter discusses the results which include the research findings while Chapter Five presents the summary, conclusions and recommendations.

1.11 Summary

The chapter dealt with the background to the study, a brief discussion of the languages investigated, statement of the problem, purpose of the study, objectives of the study, research questions, and significance of the study, theoretical framework, and definitions of operational terms, limitations of the study, scope and delimitations of the study and the structure of the thesis. The next chapter will focus on literature review.

CHAPTER TWO

LITERATURE REVIEW

2.1   Introduction

Studies have been conducted on some of the languages under investigation in this study. However only Tonga among the languages investigated has been analysed in a comparative study by Hilary (2009). The review will briefly look at some of the works, studies and publications available in the languages investigated. The challenges faced were the non availability of publications in two of the languages that were investigated as these have not been published. These are Ila and Sala.

The literature review on comparative studies includes publications in Kaonde, Lenje, Nyanja, Soli and Tonga; literature concerning Proto-Bantu since it is established that Proto-Bantu is the hypothetical ancestor language of modern Bantu languages. The researcher also reviews the different classifications of the Bantu languages by various scholars. The literature in other languages on which scholars have carried out comparative studies using lexicostatistics have also been reviewed.

2.2   Previous work on Kaonde, Lenje, Nyanja, Soli and Tonga.

Various scholars have carried out research on some of the languages investigated in this study while only few published works in Lenje have been reviewed such as Madane (1908) who published a book whose title is 'Lenje Handbook,' where he presented Lenje grammar using the English alphabet in accordance with a few necessary rules. He wrote in a plainest guide for Englishmen in getting to speak, write, and understand a Bantu dialect. The book presents extracts from native stories with translation notes with the help of Rev. Father Torrend. Other published works reviewed in Lenje is religious material that includes

the Lenje Bible whose translation took place from 1990 and was published in 2003 by the Bible Society of Zambia. Lenje is also included in Chitebeta's (2007) contrastive study of the Noun Phrase of Tonga and Lenje.

The literature reviewed on Nyanja (also known as Chewa) includes Miti (1988) who undertook a comparative study of the Zambian Chewa (Ce), Chinsenga (Ns) and Chingoni (Ng). As alluded to in 1.1.5 of chapter one in this study, these are the three varieties he referred to collectively as Zambian Chinyanja. He concentrated on the internal classification of these languages to establish how close or how different these varieties are. He put particular emphasis on their tonal relationships. His study also includes a lexicostatistical analysis and an overview of the phonology and morphology of these varieties.

Miti's study reveals that there are five varieties of Zambian Chinyanja and these are; Ce1, Ce2, Ns1, Ng and Ns2. He established that "Ns1 and Ng are tonally identical although this is not the case when lexicostatistics are considered because the focus is on the lexicon.

Apart from the research works carried out in Nyanja by different scholars cited and others not mentioned in this study, Lehman (2002) published a Nyanja textbook entitled, '*An outline of Chinyanja Grammar*' in which she discusses the sound system of Chinyanja in chapter one of the book and the morphology in chapter two. She came up with 18 noun classes and their prefixes.

Zimba (2007) carried out a comparative study entitled "*The Impact of Using Nyanja as a Language of Literacy in a Predominantly Tumbuka Speaking Area.*" He compared Nyanja and Tumbuka and established that many Tumbuka speaking pupils faced problems in

understanding Nyanja. His argument is that the difference that exist between Nyanja and Tumbuka as highlighted below makes it difficult for Tumbuka speaking pupil to comprehend the language of literacy.

Nyanja	Tumbuka	Gloss
funso	fumbo	'question'
danga	lingo	'the light'
ndio	dende	'relish'
cikwati	nthengwa	'marriage' (**Source**: Zimba 2007:143)

He further states that pupils made a lot of errors whenever they tried to read, speak or write Nyanja. His conclusion was that the difficulties that pupils faced in Lundazi contributed to the poor results they got in initial literacy lessons. Further, Zimba recommends that pupils in Lundazi rural schools should use Tumbuka for learning literacy skills in Grade One. However, the degree of relatedness between Nyanja and Tumbuka is not mentioned in Zimba's study. Therefore there is need to study the relatedness of the regional languages that are used for initial literacy skills to the languages pupils speak in particular areas. It is the purpose of this study to fill this gap.

Ngalande (2007) undertook a corpus-based study on "*A Logical Analysis of selected texts in Nyanja',* where he concentrated on the logical analysis." His research does not involve issues of inter-relatedness of Nyanja to any other language. However it is part of the previous works in Nyanjya, one of the languages under study.

Kikaonde is one of the seven Zambian languages whose orthography was approved by the Ministry of Education (1977) where Kaonde vowels, consonants, word division, conjunctions, pronouns, compound nouns, reduplication of words or stems, idiophones, adjectives, verbs, adverbs, interjections, borrowed words and place of names are discussed. The text was useful to the study especially on vowels and consonants in Kaonde. However few literary works have been published in Kikaonde and the following have been reviewed in this study such as Mikaili (1968) who published a book entitled, '*Byalwile Mushala Mema Ke Mali*' and Mutembakyalo (1988) who published a book entitled, 'Kyapusana.' Linguistic description have been carried out and few text books have been written in and on Kikaonde. For example, Wright (1985) published a glossary entitled the '*Kikaonde Vocabulary;*' Boroghall (1995) published the first '*Kikaonde Bilingual Dictionary;*' Further, Wright (2007) produced '*An Outline* of *Kikaonde Grammer.*' The text is made up of five chapters where the author discusses the sound system and the structure of Kikaonde language.

Mambwe (2008) carried out a dialectological study on 'Some Linguistic Variations of Kaonde. He established that there are major dialectal variations at the various levels of linguistic analysis. However, Mumbwa dialect displayed more variations in relation to other dialects investigated due to the long distance between the Mumbwa dialect and the other two, while Lubango-Solwezi dialects displayed minimal linguistic variations between them because of their geographical closeness. Mambwe's study established that there are more lexical differences exhibited among the dialects, syntactical and phonological variations while the three dialects shared more similarities in their morphological structure.

While literary works have been written in Kikaonde and linguistic works have been carried out on Kikaonde, no research has been carried out to study the degree of relatedness between Kikaonde, and any of the other languages under study particulary Tonga, the language which is used for initial literacy and as content subject in the Kaonde Speaking area: Mumbwa District, Central Province, Zambia. This research is intended to bridge this gap.

There is no publication in Soli grammar. Soli language has not been published by many authors. The researcher learned that the translation of the Soli Bible is not yet finalized according to Chongwe Radio Station, while few religious materials in the form of pamphlets were found in the villages around Lwiimba Primary School in Chongwe District, printed by the Jehova's Witness Church called the Watch Tower Bible and Tract Society such as '*Nyumfwilani Lesa*' published in 2011; '*Nyumfwilani Lesa: Kwamba eti Mukabe Nebuyumi Butapu*' published in 2011 and '*Makani Aina kufumina kuli Lesa*' published in 2012.

Tonga has been more thoroughly studied and written on by many authors. The literature reviewed includes:

Torrend (1931) authored the dictionary called, '*An English Vernacular Dictionary of the Bantu Botatwe Dialects of Northern Rhodesia';* while Hopgood (1940) carried out a comparative study of Tonga and Ila in his book entitled '*A Practical Introduction to Chitonga';* Collins (1962) published a textbook entitled 'Tonga Grammar, while Meeusen (1963) published an article on the '*Morphotonology of Tonga Verb*'; Carter and O'Brien (1970) compiled a 'Tonga Programmed Course in 70 lessons; Hachipola (1988) studied the

eleven indigenous languages in Zimbabwe where Tonga was included as one of the marginalised languages except for Shona and Ndebele used for educational purposes; Nkolola (1997) undertook a study on 'The Analysis of the Morphology, Syntax, Semantics and Phonology of the Applied, Causative and Passive Verb Extensions in Tonga.' Further, Carter (2002) Produced a book entitled '*An Outline of Chitonga Grammar*' which concentrated on the Sound System and Orthography; Morphology and Syntax; Chitebeta (2007) studied the '*Noun Phrase of Tonga and Lenje: A Contrastive Study*' while Hambaba (2008) examined the '*Systematic Processes of Determination and Modification in English and Tonga.*'

Hillary (2009) carried out a comparative study entitled, *'A Lexicostatistical Survey of Selected North-Eastern and South-Eastern Bantu Languages.'* The study included five Bantu languages namely Citonga, Ikalanga, Kikamba, Kiswahili and Setswana. The study concentrated on establishing the degree of relatedness between these languages. Hillary's study established that the degree of relatedness between Citonga and other languages was as follows:

Citonga vs. Setswana   38.4%

Citonga vs. Ikalanga   44.4%

Citonga vs. Kiswahili   45.6%

Citonga vs. Kikamba   43.2%

While Hilary established these percentages of the degree of inter-relatedness for Citonga, the study does not include the inter-relatedness of the Zambian languages dealt with in this study apart from Citonga.

Musale (2009) studied '*The Grammars of Compound Nouns in Tonga*, and more recently, Sibajene (2013) carried out 'A Dialectological Study of Tonga,' in which he particularly investigated the two Tonga dialects namely, Plateau Tonga and Valley Tonga. The purpose of his study was "to identify phonetic, morphological and syntactic variations. He also purported to establish the linguistic nature of lexical differences between Valley Tonga and Plateau Tonga" (page 4). His study does not include the inter-relatedness of other languages with Tonga other than the dialects of Tonga.

The reviewed literature on the Zambian languages has revealed that there still remains a gap in linguistic studies particularly the lexicostatistical study of the languages to establish the inter-relatedness of languages. It is the intension of this study to fill this linguistic gap.

2.3   Proto-Bantu (PB)

The term Proto-Bantu was coined by Malcom Guthrie referring to a hypothetical parent language considered to be the direct ancestor (or parent) language of modern Bantu languages. Proto-Bantu was reconstructed through the application of the comparative method to present day Bantu languages. According to Guthrie (1970b) Proto-Bantu was spoken somewhere in the Congo Basin. On the other hand, Postnanski (1966) says that Oliver and Foge (1962) agree with Guthrie and suggested that the earliest Bantu-speaking peoples may have been hunters and fishermen who moved along the Congo and encountered and adopted the cultivated plants of the earliest traders and migrants from the

South-east Asia. According to Greenberg (1948), the Bantu peoples and languages originated in and spread from the Cameroon/Nigeria border area.

Miti (2006) states that most scholars favour Greenberg's theory that Bantu peoples and languages originated and spread from the Cameroon/Nigeria border area.

Further, Miti (2006) cites Guthrie (1967-71) as currently the main reference work for Proto-Bantu phonology. Guthrie hypothesized the historical development of Proto-Bantu. He did this by comparing items from several modern Bantu languages. His findings were that Proto-Bantu probably developed into two dialects which he code-named PB-A and PB-B respectively. Guthrie's hypothesis was that PB-A originated in the west, whilst PB-B originated in the east. In his opinion, the two were dialects of the original ancestor language of present-day Bantu languages. He code-named this ancestor language as PB-X. As mentioned above, Guthrie's work constitutes the main reference for Proto-Bantu phonology and lexicon. He made a provisional list of his PB-X stems and radicals of common Bantu in Guthrie (1967-71). He used a star on items that are radicals and stems for hypothetically reconstructed items, lexical items which are assumed to be the Proto-forms from which all the cognate items in present-day Bantu languages are derived. Some examples of Guthrie's starred items are provided in Figure 2 below

	Proto Bantu	**English**
a.	*-ba-	be, become
b.	*-bab-	sting
c.	*-bed-	boil

d.	*-bede-	body
e.	*-dem	cultivate
f.	*-cek-	laugh
g.	*-demi	tongue
h.	*dom	bite
i.	*-jeda	path
j.	*-uku	die
k.	*-kumu	chief
l.	*-ni-	defecate
m.	*-nu-	drink
n.	*-nua	mouth
o.	*-pa-	give
p.	*-pet-	pass
q.	*-pum-	come/go out

Figure 2: Lexical items assumed to be Proto Bantu (Source: Miti 2006:65)

As shown above, it is conventional to place a dash before and after a verb radical, but a dash only before a stem. For example, the Tonga word -bili 'body' is a stem while *-lim-* 'cultivate' is a verb radical. Stems and radicals are the bases that remain when nouns and

verbal forms are stripped of all affixes. The dash placed before and/or after a starred item indicates the position from where the affixes have been removed.

The current study has utilized some of Guthrie's reconstructed Proto-Bantu forms, to try and compare the relatedness of the languages under investigation. The Proto-Bantu forms have also helped to establish whether the cognate items between the languages under investigation are related due to their descent from the Proto Bantu.

2.4   Classification of Bantu Languages.

Some scholars have tried to classify African languages, including Bantu languages. For example, Doke (1943) came up with a tentative classification of Bantu languages where he used zones, namely, (1) Central zone, (2) Congo zone, (3) East Central zone, (4) Eastern zone, (5) North-eastern zone, (6) Northern zone, (7) North-western zone, (8) South-Central zone, (9) South-Eastern zone, (10) West-Central zone, and (11) Western zone.

Guthrie (1948) argued that Doke's attempt to classify Bantu languages was inconclusive because he did not make any reference to those Bantu languages that had not been documented at the time. However, Guthrie's argument does not specify which languages and dialects Doke (1943) did not include in his classification due to lack of published materials. Guthrie included some of the languages under study, particularly those in the Bantu Botatwe group in his classification as mentioned in the background to this study. Ila is classified in Zone M, Group 60, and Cluster 63. Kaonde: Zone L, Group 40. Lenje: Zone M, Group 60 and Cluster 61. Nyanja: Zone N group 30, cluster 31. Tonga: Zone M, Group 60, and Cluster 64. As shown above, Guthrie used capital letters for assigning zones. Guthrie's work provides the basis for the classification of Bantu languages.

Miti (2006) cites Doke as one of the linguists who have significantly contributed to the study of Bantu languages. Most notable was his classification of the languages. He divided the Bantu languages into zones and groups. He described the zones as being largely geographical. However, Doke cautioned that individual members of a given zone may today be living among members of a different zone due to tribal migrations. In his classification Doke used figures for zones, for groups, and for individual languages. He used Guthrie's method by assigning small letters to indicate dialects. Nevertheless, he departed from Guthrie's method by not using capital letters for zones. Rather, he used the first two digits to represent the zones, the third digit standing for the language group (or cluster), and the fourth for the language. He placed the languages that share the same phonetic and grammatical features in the same group. Doke claimed that such languages were mutually comprehensible. Miti (2006) summed up Doke's classification as presented below:

Zone name	Zone code	Number of groups
North-western	10	6
Northern	20	8
Congo	30	7
Central	40	4
Eastern	50	11
North-eastern	51	3
East-central	52	5
South-eastern	60	5

South-central	61	1
Western	70	5
West-central	71	4

From Doke's classification of Bantu languages discussed above, the researcher observes that Doke did not classify some of the languages under investigation. These include Sala, Kaonde Ila and Soli. As mentioned earlier, the reason might be that these languages had not yet been documented at the time of Doke.

Guthrie (1948) used the practical method to compare Bantu languages. He selected one particular language as the starting point and, moving outwards, he established languages that were similar until a certain point where he noticed that the languages were becoming different, which meant he had reached a new zone. The researcher draws from Guthrie's theoretical method of beginning with one language, Sala, as a starting point then moves outward to compare it with other languages under investigation. Guthrie also applied what he called the Principal and Subsidiary criteria. He says that it was necessary to divide into two groups the criteria to be used for identifying languages as Bantu, because there are some languages in which contraction and attrition have to be postulated to such an extent that it becomes extremely difficult to apply some of the criteria.

Guthrie's Principal Criteria are listed as follows (Guthrie 1948:11):

2.4.1   Principal Criteria

(a)   A system of grammatical gender, usually at least five, with the following features:

(i) The sign of gender which is a prefix and by means of which words may be assorted into a number of classes varying roughly from ten to twenty.

43

(ii) Regular association of pairs of classes to indicate the singular and plural of the genders, there are also one-class genders where the prefix is sometimes similar to one of the plural prefixes.

(iii) An independent prefix as the sign of the class of a word. Any other word which was subordinate to it has to agree with a dependent prefix.

(b)   A vocabulary, part of which could be related by fixed rules to a set of hypothetical common roots.

The principal criteria mainly apply to the noun class system.

Guthrie (1948:11-12) also lists what he calls subsidiary criteria, as follows:

2.4.2 Subsidiary Criteria

A subsidiary criterion is a set of invariable cores, or radicals, from which almost all words are formed by an agglutinative process. These radicals have the following features:

(i) They are of the structure Consonant + Vowel + Consonant (CVC)

(ii) When a grammatical suffix is attached to the radical there is formed a 'base'

(iii) When a non-grammatical or lexical suffix is attached to the radical there is formed a 'stem' on which words identifiable as nominals are built.

(iv) A radical could be extended by an element found between it and the suffix. Such elements termed as 'extensions' are composed either of Vowel + Consonant, or of a single vowel.

(v) The only case of a radical occurring without a prefix of any kind occurs in verbals used as interjections.

The subsidiary criteria mainly applied to verbs.

A good example to explain the above subsidiary criteria is where words are formed by an agglutinative process. For example, in Tonga, the root *belek- 'work' is the radical (=root)* using the agglutinative process, we can attach affixes to this radical to come up with the word *tatukababelekeli (ta-tu-ka-ba-belek-el-i)* 'we will not work for them.' The components of this word are negative marker (NM), subject marker (SM), tense marker (TM), object marker (OM), verb root (VR), applied extention (APL) and verb ending (VE) as shown below:

ta-	tu-	ka-	ba-	belek-	el-	i
|	|	|	|	|	|	|
NM	SM	TM	OM	VR	APL	VE

This study included both nouns and verbs in the basic vocabulary of the languages under investigation. The verbs analysed are what Guthrie called stems that is root + suffix. Most of the nouns are analysed with their prefixes.

The other subsidiary criterion advanced by Guthrie (1948) is a balanced vowel system consisting of one open vowel 'a' with an equal number of back and front vowels. An example of a balanced vowel system is given in Figures Three and Four.

	FRONT		BACK	
HIGH		i		u
MID		e		o
LOW			a	

Figure 3: Balanced five vowel system in radicals consisting of one open vowel 'a' (Source: Guthrie 1948)

	FRONT		BACK	
HIGH		i		u
HIGH-MID		e		o
LOW-MID		ɛ	ı	
LOW			a	

Figure 4: Balanced 7 vowel system consisting of one open vowel 'a' (Source: Guthrie 1948)

Using these criteria, Guthrie (1948) classified Bantu languages into Zones, then groups and lastly into sub-groups or clusters. Zones were denoted by capital letters, groups by numbers and sub-groups by small letters. He had zones A to T, groups 10 to 80 and subgroups 'a' to 'd' depending on the number of languages in the group.

Guthrie's 1948 classification of Bantu languages has become the major reference work for linguistic scholars in Bantu languages. Some of the languages under investigation have been placed in group 60 in Guthrie's classification as follows: Zambian Tonga is classified as M64a that means the language is found in Zone M, belonging to group 60, and is the fourth language in the group which is the first dialect (M64a). Group 60 consists of four

languages: ciLenje (M61), ciSoli (M62), ciIla (M63), and ciTonga (M64a). According to Guthrie (1967:81) "the Zambian Tonga has the following three dialects: ciTonga (M64a) Toka (M64b), and Leya (M64c)." The present study concentrates on some of the languages that Guthrie classified in zone M group 60 particularly those that belong to Bantu Botatwe group including Sala which is not classified by Guthrie and some of those languages that relate to them such as Kaonde (L40), Kaonde-Ila (not classified by Guthrie) and Nyanja (N31).

Cole (1969) recognized Greenberg's authorship of "**The Classification of African Languages"**, published in 1948 and his series of articles in the Southwestern Journal of Anthropology entitled "**Studies in African Linguistics**." Greenberg's (1963) classification of African languages followed the genetic type of classification, based on three principles:

a) Regular morpho-semantic similarities, which is concerned with the comparison of resemblances involving both sound and meaning in specific forms.

b) The principle of mass comparison that means he compared a large number of lexical items from many languages of the African continent.

c) Linguistic correspondences where the word meaning is quantifiable as much as it is uniquely translated in some language or set of languages. He objected to the previous attempts based on typological or ill-defined structural criteria. Greenberg argued that structural features alone should not be used in determining genetic relationships between languages.

The current study has borrowed from one of Greenberg's principles, namely the principle of mass comparison. The study compares 200 lexical items from eight Zambian languages

namely, Ila, Kaonde, Kaonde-Ila, Lenje, Nyanja, Sala, Soli and Tonga. According to Campbell (2004) mass comparison aims at simply showing which languages are related to each other. Further, Greenberg (2001) suggested that the mass comparison method is useful for preliminary grouping of languages known to be related as a first step towards a more in-depth comparative analysis.

Greenberg (1963) classified African languages into the following major-families: Congo-Kodofanian, Nilo-Saharan, Afro-Asiatic, and Khoisan. In respect to this study, the Congo-Kodofanian is the most relevant of the four families because Bantu languages belong to this family. Greenberg divided the Congo-Kodofanian family into two main branches namely, Niger-Congo and Kodofanian and coded them IA and IB respectively. Greenberg rejected the typological classification of languages, on the grounds that unrelated languages in remotely separated parts of the world have been known to evolve similar case systems or gender systems independently of each other. Unrelated languages may be categorized as isolating, or agglutinating, or inflectional in structure. Languages categorized as having case, or gender systems may lose these systems in the course of time and languages without the systems may acquire them. However, Greenberg did not include borrowings and chance resemblances in his classification. For this he received criticism from the Africanists based in Europe. For example Westphal accused Greenberg of using the term family in a wider sense than it is usually given (Cole 1969).

Miti (2006:38) states that "the major controversy in regard to Greenberg's classification stemmed from disagreement and misunderstanding of definition of certain crucial terms." In 1957 Westphal attacked Greenberg for basing his classification mainly on lexical criterion rather than on morphological factors. He further criticized Greenberg for dealing

with languages which belong to two different language types as if they were one family. Despite the criticisms by Westphal and others, comparative linguists to this day generally appreciate Greenberg's pioneer work on the classification of the languages of Africa (Miti 2006:40).

Typological classification is a classification based on structural types across languages. According to Masayoshi and Thoedora (1996:92), typology is "the study of linguistic patterns that are found cross-linguistically; in particular, patterns that can be discovered solely by cross-linguistic comparison. In the nineteenth century there arose a classification of languages based on the morphological structure of words commonly called Morphological Typology."

Matthias (2003:33) indicates that "languages are divided into three types namely: Isolation, Agglutinative and Flectional."

a) Isolating languages such as "Chinese and Africaan have very simple forms in which grammatical relationships are not expressed in the word"

b) Agglutinative languages, for example Bantu, Turkish, Finish and Magyar have "words consists of several distinguishable morphemes, each indicating distinct grammatical category."

c) Flectional languages: Indo-Germanic and Semitic families. "Words involve complex internal morphological changes, so that morpheme boundaries are obscured or lost."

The languages investigated belong to the agglutinative type. According to Bodmer and Lancelot (1972), the term 'agglutinative' was introduced by Wilhlem von Humboldt in

1836 to classify languages from a morphological point of view. The word is derived from the Latin verb *agglutinare*, meaning 'to glue together.' In agglutinative languages, each affix typically represents one unit of meaning such as 'diminutive,' 'past tense,' 'plural,' and others. Bound morphemes are expressed by affixes. Affixes do not become fused with others and do not change form conditioned by others. For example in Tonga *-ul-* 'buy' is the radical (=root) using the agglutinative process we can add affixes to this root to come up with the word *ta-ndi-ka-ba-ul-il-i* 'I will not buy for them' which has these morphemes:

ta-	ndi-	ka-	ba-	ul-	il-	i
|	|	|	|	|	|	|
NM	SM	TM	OM	VR	APL	VE

2.5   Some Comparative Studies and Lexicostatistics

Mildred (2007) studied the main Indo-European language groups using lexicostatistical analysis. She hypothesized that Italic is related to Hellenic. She also discovered the phonological and semantic similarities of Latin and Greek that had a degree of relatedness of 69%. The following examples are some of the Latin and Greek words that are related phonologically and semantically:

Latin	Greek
duo	du': o (dual)
ego	ego': (ego)
pes	pu': s (foot)
humus	xamos (ground) xamai (on the ground)

folium	fu'llon (leaf)
frater	phrater (brother)
petra	p'etros (stone)
domus	do'mos (home)
rivus	rheos (river)

The first two examples in Latin and Greek: *duo* and *ego* respectively, are cognates both phonologically and semantically. The Latin word *frater* and Greek word *phrater* are considered to be cognates although the Latin phoneme *f* is changed into Greek phoneme *ph-* which has the sound as *f.* The phonetic transcription in both Latin and Greek is a labial dental fricative [f].

Mildred (2007) cites Dyen (1991) and Ringe (2005) who surveyed the lexical innovations of 46 lexeme list of Baltic and Slavic languages and established a relatedness of 60%. Few examples of vocabulary from Baltic and Slavic are given below:

Baltic	Slavic	Gloss
ranko	ranka	hand, arm
nage	noga	foot, leg
zwaigstan	zhvaigzhde	star
zirgis	zhirgas	horse.

The study showed that close genetic proximity of both groups is evident to anyone familiar with any two Baltic and Slavic languages. Some selected words and phrases may not even require translation. For example, the Baltic word *ranko* and the Slavic word *ranka* 'hand, arm' differ only in the ending vowels /o/ and /a/ respectively. This makes it easy to identify cognates.

Lexicostatistics has been scarcely carried out in languages of Africa. For example, the Mande Languages form a coherent linguistic unit in West Africa. They number 30 to 40 distinct languages, or dialect clusters, and are spread from eastern Senegal to western Burkina Faso, scattered groups being found still further eastward. Their speakers constitute the majority of the population in Gambia and Mali and form important groups in several other West African countries.

According to Hans (1972), the Mande Languages were classified wrongly, among the Nigritic languages which cover the largest part of Africa south of the Sahara. In fact, this view has already been abandoned by a part of Africanists. Others cling to the idea that Mande has seperated first, before any other linguistic entity, from the Common Nigritic stock. In any case the Mande Languages are fairly different, in lexis as in structure, from the 'Nigritic' Languages of West Africa.

Comparative studies of the Mande Languages are not yet sufficiently advanced. The results of a lexicostatistical test undertaken convinced Hans (1966) that Mande must be related to Hamito-Semitic, and that it is closer to Cushitic und Chadic than to its northern branches Berber and Semitic. In a later study Hans (1972) demonstrated that the common numeral system of the Mande languages strongly resembles that of Central Cushitic.

To come up with his conclusions, Hans (1972) used the wellknown 100 diagnostics items test list proposed by Maurice Swadesh. He noted that Fula (or Peul in French terminology) is a very widely spread West African language, spoken by several million people from the Senegal River to the East of Lake Chad. It is not related to Mande on a recognizable level, and its alleged relations to other languages of West Africa are uncertain. It forms a closer linguistic unit only with two other Senegal languages, Wolof and Serer.

The count of the common basic vocabulary of Fula and Mande was between 13-14 percent. Therefore, Hans (1972) concluded that Fula is a descendant of another linguistic group that once existed in the Saharan area, which he called Mauritanian.

David and Calvin (1977) undertook a study on Comparative Reconstruction of Proto Jicaque- Subtiaba- Tequislateco and they discovered that these languages were not classified because the affiliations of the Xinca, Lenca, Jicaque and Paya languages were so uncertain and controversial that Mason (1940) felt that it was best to leave these languages unclassified.

In 1953, Greenburg and Swadesh investigated the structure and vocabulary of Hokan and Jicaque languages and stated that "we find unmistakable evidence that Jicaque is a Hokan language and that it is related to Hokan-Coahutecan is clearly evident from the quality and quantity of agreements." (1953:216, 220).

Using a lexicostatistical study, it has been concluded that the language closely inter-related to Jicaque is Tequislateco. For example, the percentages of inter-relatedness established were:

Supernec to Jicaque 23 percent (closest).

Tequistlateco to Jicaque 13 percent (more remote).

Another study was carried out by Rensch (1966) who states that there is a close relationship between Subtiaba and Tlappanec. This stand is supported by Sapir (1925:403) who confirms that "Subtiaba and Tlappanec are really only dialects of a single language…" It is believed that these two languages are mutually intelligible or nearly so.

The review of Oltrogge and Rensch's (1977) study helped the researcher in seeing how other linguists have applied the lexicostatistical method in comparative studies to come up with the degree of inter-relatedness of languages.

While Doke (1945) studied Swahili and its closest relatives, the biography of Nurse and Hinnebusch (1993) made clear the knowledge of Swahili and its dialect Sabaki.

Walsh (2000) made reference to Nurse who examined the Bantu sub-grouping using both lexical and phonological evidence. He divided his 'Corridor into three:

West Corridor (MWIKA) includes Pimbwe, Rungwa, Fipa, Rungu, Wanda and Namwanga.

Central Corridor (NYIKA) includes Lambya, Malila, Nyika and Safwa.

East Corridor (NYAKYUSA-NDALI) includes Nyakyusa and Ndali.

Nurse's classification of the three groups above was echoed by Labraussi's (1998 and 1999) research findings, who also recognized three groups in the corridor as Mwika, Nyika and Nyakusa which includes Ngonde and Ndali.

Using a lexicostatistical classification, based on data from 14 Corridor languages, Fourshey (1996:6) returned to the view that "all the languages of the Corridor derive from a common ancestor" she calls 'Proto Rukwa.'She divided Proto-Rukwa into Proto-Rungwe and Proto-Mbozi, the former is equivalent to Nurses's 'Nyakyusa-Ndali and the latter to his 'Mwika-Nyika.'

Recent Comparative Bantu studies have come up with notable results particularly in phonology, morphology and the lexicon. Miti, (1996) carried out a comparative study involving vocabulary correspondences between Tshivenda and Northern Sotho and also between Tshivenda and other Bantu languages spoken in the Southern African Development Community (SADC) region such as Cinyanja, Cishona, Citumbuka, Xitsonga and Isizulu. He used a lexicostatistical analysis to illustrate the extent to which such an analysis may conceal, rather than reveal, similarities between language varieties. Among the languages that Miti (1996) studied, he established that the most closely related languages were Tshivenda and Shona; and Chinyanja and Citumbuka. Each of these two pairs of languages shares about 42 percent of core vocabulary.

Further, Miti (1996) proposes that the criteria for recognizing cognates should be more inclusive so as to treat all those items that can be correctly interpreted by speakers of other related languages without exposure to the languages. He concluded that many languages of Africa are related to varying degrees. Some are so similar in their vocabulary that they are mutually intelligible; others have differences in the sound segments that constitute lexical items. He established that Tshivenda and Xitsonga share 36% while Xitsonga and Zulu share 27%. Miti discovered that the least relationship is between Northern Sotho and Tumbuka which is at 12% of the basic vocabulary. Miti (1996:173) concluded his studies

by saying that "since all knowledge is transmitted through language, it is crucial that similarities between the various related African languages be highlighted. In that way, Africans will be better able to share the very rich and varied indigenous knowledge that exists on the continent." This study drew insights from Miti's (1996) criteria for determining cognates, which is more inclusive to try and determine the degree of relatedness between the languages under study.

Other scholars have carried out studies of Bantu languages using the lexicostatistical method to analyze data in order to establish the degree of relatedness of a particular group of languages. Batibo (1997) is one of the scholars who based his studies on the lexicostatistical analysis of nine Botswana Bantu languages namely, Setswana, Setswapong, Sebirwa, Sekgalagati, Ikalanga, Thimbukushu, Shiyeyi, Chisubiya and Otjiherero. In his study, Batibo discovered a very high degree of inter-relatedness between Setswana, Sebirwa, Setswapong and Sekgalagadi which he rated at 81.8 percent. He rated the degree of inter-relatedness between Shiyeyi and Chisubiya at 64 percent while Shiyeyi and Otjiherero are rated at 44 percent. Thimbukushu shares only 51.5 percent with Shiyeyi and Chisubiya. This indicated reasonable high intercomprehension between the selected languages. Batibo went further to confirm the degree to which the languages retained the proto-Bantu and Eastern Bantu vocabulary. He observed that Chisubiya and Shiyeyi have the highest percentages of retained Proto-Bantu/Eastern Bantu vocabulary. This is due to their proximity to the North Eastern Bantu nucleus as implied by Ethret (1998).

Batibo (ibid) further observed that the lexicostatistical method is more hierarchically refined than those of the comparative method. He also discovered that lexicostatistics provides more precise information about the degree of inter-relatedness between dialects.

The study used a carefully selected list of 250 basic vocabulary items from common proto-Bantu and Eastern Bantu.

Following Batibo's (1997) study, the researcher undertook a comparative study of seven Zambian languages, in which the percentages of inter-relatedness of these languages were studied using the lexicostatistical approach, which involved the quantitative comparison of lexical cognates.

Mildred (2007) compared the main Indo-European language groups where she used the classical method by considering the actual phonology of two sets of opposite cognates. She states that the facts demand to take notice of the actual phonological similarity of lexico pairs and compare words in a phoneme-by-phoneme, rather than word-by-word fashion. For example,

English	**German**	Canada
two /tu: /	zwei /tsvai/	yereki'
three /Ɵri: /	drei /drai/	yerku
four /fo: r/	fier /fi: r/	chors

Using the classic method, Mildred (2007) compared phoneme-by-phoneme to identify the cognates in the above listed words. She affirms that /tu: / and /tsvai/ look related as opposed to *yereki'* because of the phoneme /t/ found in both items. Equally /fo: r/ and /fi: r/ were considered cognates because of the presence of the phoneme /f/ in both items.

Further, she presents examples of cognates for Latin and Greek as follows:

Latin	Greek	Gloss
pes	pu: s	'foot'
humus	xamos, xamai	'ground', on the ground
folium	fullon	'leaf'
petra	petros	'stone'
domus	domos	'home'
rivus	rheos	'river'

Mildred (2007) established that the degree of inter-relatedness between Latin and Greek is at 69 percent. In all the examples given above from Mildred's study, the researcher observes close phonological and semantic proximity that can be explained by assuming genetic unity of Greek and Latin languages.

In Hilary's (2009) study, the two categories of cognates in the languages studied were established as follows: (a) Cognates that have been retained from Proto-Bantu and (b) Cognates shared among three or four of the languages examined but are not inter-relate with Proto-Bantu. The study established that most of the cognates retained from Proto-Bantu are numerals ranging from 2-5, a few names of domesticated and wild animals, and body parts. The majority of retained cognates are verbs. She also established that Kiswahili retained more Proto-Bantu forms than any other language under investigation. Hillary (2009) identified several sound changes between Proto-Bantu and the languages under study. These include denazalization, deletion, palatalization, lateralization, fricativization and voicing. She further noted that Ikalanga was found to have borrowed some sounds and

also lexical items from Setswana due to Ikalanga's long existence in the same area and also due to Setswana's status in Botswana. On the other hand Setswana was found to have unique sounds that are not found in any of the other languages under investigation.

Hillary (2009) cites Heine (1974) who compared the Kalenjin languages of Kenya following a 200 word list. Heine (1974) asserts that the lexicostatistical method was useful in establishing genetic relationships. She also refers to Hinnebusch (1999) who indicates that Greenberg's work used a statistical method in nature referred to as mass lexical comparison, while Guthrie devised his own statistical approach in his Comparative Bantu where he developed several measures: a coefficient of Bantuness, Indices of relationship and modulus of dispersion.

In a paper presented at the Swadesh Centenary Conference, Grant (2009) asserted that the best approach to linguistic data for lexicostatistics is that of Don Ringe and his associates (*Ringe*, Warnow, Taylor and Warnow 2002) who take a taxonomic approach to this matter. They see entries on lists as characters, and the forms in each language denoting such entries as states of characters. Once the data have been assembled and recurrent sound correspondences have been identified and recognized, they need to be assigned codes for the purpose of quick reference, for example, *aaa* meaning all three languages for the items in question are correct; *aab-* means that the forms in the first two languages are cognates, but the last language's form is not; a*ba-* showing that language one and three are cognates but not in two. This coding procedure is the same in principle as that used in normal lexicostatistical anaysis. This study used the same principle of coding procedure as that of Batibo (1997), Miti (1996) and Hilary (2009), that uses the + code for cognates and a − code for non-cognates. For example, +++ means the items in all the three languages are

cognates. -++ means the first language is non cognate while the ones in the last two are cognates. However, this study used two additional codes such as a hash (#) sign to represent a second set of cognates for example, -++++-##- meaning that the second, third and fourth languages are first cognates while the sixth and seventh languages are second set of cognates. A caret (ˆ) sign was used to represent a third set of cognates. For example, ++-###ˆˆthe first two languages are the first set of cognates. The fourth, fifth and sixth languages are the second set of cognates and the seventh and eighth languages are the third set of cognates. Four different types of codes were used in order to analyze the data adequately.

This current study utilized the lexicostatistical approach and a modified vocabulary of Miti's (1996) 200 word list of basic vocabulary but focused on seven Zambian languages that have not been explored by previous scholars in relation to some of the languages of the Bantu Botatwe group.

2.6 Summary

The present chapter has reviewed research works and publications in Kaonde, Lenje, Nyanja and Tonga. The discussions in this chapter also included the classification of Bantu languages, some comparative studies and lexicostatistics. The next chapter will deal with methodology.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

This chapter presents the methodology that was applied in the study. The approach was both qualitative and quantitative. The study location reveals the different parts of the country that were covered in the present study. The study population shows the targeted population who are native speakers of the languages being examined. The chapter also discusses sampling techniques in which purposive sampling technique was used. Data collection techniques are also discussed in this chapter. The instruments for data collection were the Swadesh list that is a basic vocabulary is discussed. The chapter also includes ethical considerations and ends in a conclusion.

3.2 Research Design

The study used qualitative and quantitative research design. Qualitative research is generally a research that utilizes interviews to explore and understand the attitudes, opinions, feelings, and behavior of individuals or a group of individuals. According to Mugenda and Mugenda (1999:155) "qualitative research includes designs, techniques and measures that do not produce discrete numerical data", while White (2000:24-25) states that "qualitative research is a stable form of research of the social sciences, politics and economics, all subjects linked with business. It is a descriptive, non-numerical way to collect and interpret information." With this approach, descriptive data is collected.

On the other hand, quantitative research adopts a natural science model of research, which searches for causes through methods such as questionnaires, inventories and demography that produce data acquiescent to statistical analysis.

Vanderstoep and Johnstone (2009) have described the quantitative design as a research that specifies numerical assignment to the phenomena under study, whereas the qualitative design produces narrative or textual descriptions of the phenomena under study. The two scholars indicate the advantages and disadvantages of quantitative research. The advantage of quantitative research is that the findings from the sample under study will more accurately reflect the overall population from which the sample was drawn. The disadvantage of the quantitative approach is that because the study contains so many participants, the answers research participants give do not have much depth. They have to be superficial, or else the researchers would be overwhelmed by information that cannot be adequately analyzed. To sum up the foregoing information, Vanderstoep and Johnstone (2009) summarized it in Table 2 below:

Table 2: Quantitative and Qualitative Research (Source: Vanderstoep and Johnstone (2009)

Characteristic	Quantitative Research	Qualitative Research
Type of Data	Phenomena are described numerically	Phenomena are described in a narrative fashion
Analysis	Descriptive and inferential statistics	Identification of major themes
Scope of inquiry	Specific questions or hypothesis	Broad thematic concerns
Primary advantage	Large sample, statistical validity, accurately reflects the opulation	Rich, in-depth, narrative description of sample
Primary disadvantage	Superficial understanding of participants' thoughts and feelings	Small sample, not generalizable to the population at large

This study applied both qualitative and quantitative approaches. It was qualitative because the researcher used narrative or textual descriptions of the phenomenon under study. The inventory of basic vocabulary of the eight languages which was used for the analysis to establish their degree of relatedness was verified by native speakers of each language through face to face interviews, for translation, verification and clarification of the data. The study was also quantitative in the sense that it applied some statistical figures such as percentages of inter-relatedness for the seven languages under study for the purpose of conclusions and interpretations. According to Bynon (1977), the Lexicostatistical method is a mathematical or statistical study of basic vocabulary of two or more languages. This study describes the vocabulary of the seven languages in a systematic way to clearly show the lexicostatistical percentages.

The researcher employed both qualitative and quantitative methods because these methods "complement each other in that qualitative methods provide the in-depth explanations while quantitative methods provide the hard data needed to meet the required objectives…" (Mugenda and Mugenda 1999:156).

3.3   Study Location

The research was undertaken in the following areas within Zambia namely; Chibwalu Village, Chongwe; City Market, Lusaka; Rusangu University, Monze; University of Zambia, Lusaka; Zambia Union Conference of Seventh-day Adventists, Lusaka; Chief Mumba's chiefdom and the surrounding villages, Mumbwa; Situmbeko Village, Mumbwa; and Chief Shakumbila's Palace, Mumbwa and in the surrounding villages. It was also largely desk research as dictionaries, grammars and other written materials on the languages under study were used to check the vocabulary items.

3.4   Study Population

Generally, population refers to people living in a given area, city, country and so on and so forth. In research population means the target group of people appropriate for a particular study. For this study, the population is divided into two: firstly, the researcher targeted the native speakers of Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga, from different parts of the country. In particular, these people were considered to have competence in their respective languages. Secondly, the researcher targeted the local language learners: 114 respondents in Chongwe District and 157 respondents in Mumbwa District. These respondents were pupils in the age range of 13 – 19 years.

3.5   Sampling Techniques

In research, sampling techniques refer to the selection of a number of people from a larger group from which information is gathered. Rudestain and Newton (1992) assert that a sample is a subset of the population that is taken to be a representation of the entire population.

There are many types of sampling techniques such as, probability sampling, which employs strict probability rules in the selection process, where every unit of the population has an equal, accurate, and non-zero probability to be selected in the sample. This method offers a high degree of representativeness. However, the method is expensive, time-consuming and relatively complicated since it requires a large sample size, and the units selected are usually widely scattered. On the other hand, "in nonprobability sampling, the chances of selecting any case are not known because cases are nonrandomly selected" (Singleton and Straits 1999:141).

Stratified Random sampling is a special form of sampling in which "the population is divided into a number of strata and a sample is drawn from each stratum. These sub samples make up the final sample of the study" (Singleton and Straits (1999: 152).

Accidental sampling is a type of non probability sampling in which the population selected is easily accessible to the researcher. Paul Vogt (2005) employed qualitative research and in other studies where representativeness is not an issue. It is also known as incidental sampling, chunk sampling, grab sampling and haphazard sampling. When this sampling technique is employed, all units for study that the researcher accidentally comes in contact with during a certain period of time are considered. For example, the researcher may stand at the University main entrance and interview a certain number of persons passing by between 11:00 hours to 12:00 hours on certain days of the week. Such samples are easy to construct and evaluate.

Singleton and Strait (1999:158) view purposive sampling as "a type of sampling where the investigator relies on his or her expert judgment to select units that are representative or typical of the population." In this study, the researcher employed purposive sampling, which uses a technique, where the researcher purposely chooses people, who are thought to be relevant to the research topic. Therefore, the researcher purposively chose to interview competent native speakers of the languages under investigation because they were relevant to the research topic. The researcher targeted both males and females for the translation of the Swadesh list from English to the languages studied.

Snowball sampling is a type of referral sampling which uses a process of chain referral: When members of the target population are located, they are asked to provide names and

addresses of other members of the target population, who are then contacted and asked to name others, and so on and so forth. Sudman and Kalton (1986) used snowball sampling to create sampling frames and are sometimes associated with probability sampling (Goodman, 1961). The researcher utilized snowball sampling because it is a method that uses initial subjects where the desired characteristics are identified using the purposive sampling technique. The researcher identified the few subjects, who named others that they knew had the required characteristics until the researcher got the number of cases required.

According to *http://www.wiki.answers.com,* snowball sampling is a non-random sampling technique where the initial respondents to a survey recruit others for a survey. Linear snowball sampling is when the first respondent recruits one more person for the survey, who in turn recruits one more person for the survey. In this study, the number of people interviewed increased in a linear manner. The related links provided good discussion of this form of surveying. Snowball sampling is the non-representative sample, but it was an effective means of surveying people with common traits who would otherwise have been difficult to find. The advantage of snowball sampling is that it is easy to implement.

3.6   Data Gathering Techniques

Different techniques of collecting data have been employed by researchers. These include "experiments, use of available data, surveys, observations, interviews and so on and so forth," (Hilary, R. 2009:15). The current study used the available research tool, which is a modified Swadesh word list, as a data collection technique (see Appendix A). As stated in the definitions of operational terms, the Swadesh word list is a basic vocabulary word list of between 100 to 1000 words utilized in the lexicostatistical method. For this study the

Swadesh list was modified to suit the African situation since in its original form it contains vocabulary items which are not all universal in the African setting such as 'snow'. While the word 'snow' may be translated into the languages under investigation, the concept of 'snow' is not known to many African countries including Zambia, due to the existing weather conditions. Therefore, the modified Swadesh list bears the universal basic vocabulary for Africa and it was translated into the seven Zambian languages being explored by the native speakers.

Nkolola M. W. (1997:6) supports the use of native speaker's intuition in collection of data. She states that, "it is generally accepted that the data of linguistic enquiry should ideally be collected from a native speaker…" The researcher is a native speaker of Tonga. However, she collected the data from other native speakers of Tonga and verified the data using her intuition to circumvent any unconscious preconception of using the researchers' dialect. The official Tonga orthography was also checked from Carter (2000) and Carter (2002) to authenticate the data.

Nkolola M. W. (ibid) also cites Atkinson et al (1982:38) who assert that if "the linguist is a native speaker of the language he is investigating, he will be able to distinguish between well-formed strings of words…" They observe that a linguist who is a native speaker of a language under investigation "is entitled to invent sentences and non-sentences to formulate and test his hypothesis." They call such abilities of a speaker of a language as linguistic intuitions; they say that linguistic intuitions "form an essential part of the database of a Chomskyan approach to linguistics which will contain not only utterances but judgment about such utterances."

To determine the degree of inter-relatedness, the researcher employed the lexicostatistical method due to its use of a limited set of vocabulary items to come up with results in numerical form.

In a situation where a survey is carried out for census or voting, information is collected using standardized questions so that every individual surveyed responds to exactly the same questions. This was applicable in this study, because a similar set of vocabulary was used in all the eight languages. The results are presented in summary using statistical tables and graphs. The study also used the available literature, grammars and dictionaries to verify the vocabulary items. The data for performance results were collected using composite and comprehension tests. (See Appendix H)

3.7   Data Interpretation and Analysis

Data interpretation and analysis was carried out using lexicostatistics. The lexicostatistical method was employed to show the percentages of lexical inter-relatedness among the eight languages under study. Descriptive statistics such as frequencies and percentages have been used as well. The data were analyzed through tables and graphs and group average trees. The data collected were transcribed to identify the sounds used in the lexical items in the different languages. Coding was applied to facilitate data analysis. Further, the performance results were presented using tables and pie charts in chapter five.

3.8   Ethical Consideration

Churchill (1995) defines ethics as moral principles and values that govern the way an individual or group conducts its activities. In view of Churchill's definition of ethics, the researcher considers Research Ethics (RE) as the application of general ethical ideas to

research behavior. According to Taylor and Bogdan (1984), ethical considerations bring about many dilemmas for the researcher or field worker. This involves clear understanding and assurance that the researcher will not contravene the informants' confidentiality or expose them to harm.

According to the Belmont Report (1979) researchers must be concerned with three ethical issues:

(a) Respect for persons: Researchers must recognize research participants as autonomous agents, and those who have diminished autonomy (for example the young, the disabled) must be granted protection. The study involved native speakers of the languages under investigations who were believed to have the competence in the languages being explored. The researcher also involved the young (pupils) people who were taking Zambian Language classes as they were relevant to the study. The researcher did not discriminate against the physically challenged native speakers of the languages under review. One person was physically challenged among those that were identified to assist in the collection of data. By involving one of the physically challenged native speakers, the researcher demonstrated that she did not exclude participants who met the condition above regardless of their physical challenges.

(b) Beneficence- The Researcher must secure the well-being of the participants, and further, maximize possible benefits and minimize possible risks. The benefits for the research participants were taken into considerations depending on how much assistance they gave to the study. In terms of risks, this type of study did not have possible risks that would have affected the research participants.

(c) Justice- There must be fairness in the distribution of benefits and possible risks across all research participants. While there were no possible risks incurred by research participants, the researcher fairly distributed the benefits (snacks and refreshments) among the research participants. The researcher also intends to share the findings with the respective communities for their knowledge and benefit.

Some of the foregoing ethical issues might not have arisen since the issue under investigation had no harm. The researcher did not force any participant, but voluntary participation was sought and permission to use people's names was granted for those that have been mentioned in this study. As suggested by White (2000) and Bunns (2000), the researcher explained the nature and purpose of the research to the informants. Identified participants were willing to participate.

During the second phase of the research, the researcher obtained letters from the Head of the Department of Literature and Languages, University of Zambia, seeking authorization for each of the districts involved in the research. A letter of introduction from the employer (Rusangu University) Vice Chancellor was obtained also. (See Appendix G)

The District Education Board Secretary (DEBS) for each district authorized the researcher to go ahead and visit the schools within their jurisdiction, purposively sampled. When the researcher arrived at each school, permission was sought from school authorities and presented the letter with the endorsement of the DEBS. The researcher then briefed the school authorities on the values and procedures that were to be used. She assured the authorities that participation by staff and pupils were voluntary. As a way of maintaining confidentiality, the pupils' names were assigned serial numbers. The names of the schools

were included in the research with full knowledge of the school authorities. Some names of the respondents for oral interviews were included with their consent.

3.9 Summary

The foregoing chapter described the methodology utilised in the study: qualitative and quantitative research design, study location, study population, data gathering techniques, data interpretation and analysis, ethical consideration. The succeeding chapter will concentrate on discussion of the findings.

CHAPTER FOUR

DATA PRESENTATION AND DISCUSSIONS OF THE RESULTS

4.1   Introduction

This chapter focuses on the data presentation and discussions of the results of the research carried out for this study. Initially, the data presentation could have been presented in Chapter four on its own, before the discussion of the results/findings. Instead, the data for the first three objectives have been presented in Appendices 'A' and 'B' due to the layout in landscape which could not fit well into the main document in portrait. (See Appendices A and B). Therefore, the criteria used for determining cognates and how the percentages of inter-relatedness were established are presented in this chapter as part of the discussions. The data presentation and discussions of the results for the forth objective is presented withing this chapter

The chapter discusses the analysis of the 200 word list used in the study in order to establish the degree of inter-relatedness amongst the languages under investigation. These languages include Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga. The results/findings, discussions and analysis are based on the following specific objectives of the study as mentioned in chapter one: (a) establishment of the degree of inter-relatedness between Sala and other Bantu Botatwe languages namely, Ila, Lenje, Soli and Tonga. (b) Ascertaining how inter-related Soli is to Nyanja, since Nyanja is used for initial literacy and taught as a school subject to Soli children in schools. (c) Establishment of the inter-relatedness of Kaonde and Tonga, since Tonga is used for intial literacy and taught as a school subject to Kaonde children in schools. (d) To ascertain whether there is any retention of the Proto-Bantu vocabulary in the languages under investigation.

Further, the degree of inter-relatedness among the languages being explored, the group averages and the hierarchy of relationship has been discussed. Also, the cognate retention from Proto-Bantu and shared cognates, as well as sound changes, and the historical implications have been discussed. The chapter closes with a summary of findings.

4.2   Criteria Used for Determining Cognates

Previous scholars have used different criteria to determine cognates. According to Swadesh (1955:124-126), "cognates are established from the word lists." Swadesh suggested a 200 word list of basic vocabulary, which he used in his lexicostatistical studies. For Swadesh (1955:124-126) the relatedness of any group of languages is supposed to be approximately comparable to the number of cognate words present in the words making up the list. The major reason for using the specific basic vocabulary lists instead of random words is that the basic vocabulary learnt in childhood is likely to change slowly over a period of time.

The Swadesh list comprises names of body parts, names of some domestic and wild animals, simple verbs and nouns for everyday activities. Swadesh (1955) also suggested different guidelines to determine cognates such as universality, culture terms and duplication. He believes that items should be universal that is, having characteristics common to all languages. For example, some items like 'ice', 'freeze', 'sea' and 'snow' may have some correspondents in African languages, but the same concepts are alien to some of the African societies that may not have such items. For example, since it is not common to have snow in Zambia, that fact alone makes the concept alien.

In Tonga 'ice' is translated as *caanda*. However, the researcher has observed that the concept still remains alien because the word *caanda* is not commonly used since the item it

refers to is not usually found in Tonga society. The word 'freeze' is translated as *kwaanda* in Tonga, which commonly affects the plants in the garden, fields or the bush, but does not carry the concept of 'freeze' especially for frozen food in the fridge, which is well known to people using freezers. Therefore the concept of *kwaanda* still remains alien. Similarly, in areas where only rivers and dams are common, the concept of sea remains alien.

Thus, Swadesh (1955:24) suggests that "culture-bound terms should be excluded due to their retention of cultural elements, a factor that makes them too closely correlated with changes in the cultural situation to serve as an index of the passage of time." He also suggests that in terms of duplication, only one word ought to be used where a concept is expressed in different ways. For example, man/male is expressed by one term in this study. Nevertheless, in the languages under study, the concepts 'man', 'male' and 'gentleman' are expressed by different terms. In Tonga, 'man' is represented by two terms such as *mulombwana*, and *mwalumi* depending on the Tonga Dialect one is using. For those using Valley Tonga, the term *mwalumi* is applied, while those using Plateau Tonga, the term *mulombwana* is applied. The word woman/female is expressed by two different terms. These are *mwanakazi* and *mukaintu.* The two terms are dialectal. Valley Tonga uses *mwanakazi* while Plateau Tonga uses *mukaintu.*

Therefore, following Swadesh's suggestion on duplication, this study, has, from the two lexical items (*mulombwana,* and *mwaalumi*) referring to 'man', only used the term *mulombwana.* Similarly, the word *mukaintu* has been used to represent the two lexical items for woman/female. The study has used the two words; *mukaintu* and *mulombwana* that come from the same dialect, which is Plateau Tonga. M*u-* refers to a prefix for singular, and *ba- in bakaintu* marks the plural form. For example, the word *mukaintu*

wangu 'my wife', *bakaintu bangu* 'my wives.' *Ba-* may also be used to denote respect and not the plural form. Thus *bakaintu bangu* is still 'my wife' but denoting respect. Throughout the study, Plateau Tonga has been used because the researcher is a Plateau Tonga native speaker and for consistency.

On the other hand, Gudschinsky (1956) suggested that any pair of equivalent morphemes may be registered as probable cognates if a minimum of three pairs of comparable phonemes or phoneme clusters are found to agree. For example, in this study, the word for cloud exhibits the following numbers of phoneme: In Tonga it is *kumbi* 'cloud'. In Ila, 'cloud' is *nkumbi*, in Kikaonde, it is *jikumbi*, in Kaonde Ila, it is called *kumbi*. However, these have been considered as cognates in the mentioned languages due to their phoneme clusters (-mb) that are found to agree in the languages. Another example is the vocabulary item mu-nyama 'animal'. It has phoneme clusters /-nyama/ which amount to a stem in all the languages under study. The item qualifies to be cognate due to this. There is also a phoneme cluster in the vocabulary item *munwe* 'finger.' It has phoneme clusters /-nwe/. In Ila, Kaonde, Lenje, Sala and Tonga it is called *munwe.* Hence the item fits to be cognate.The vocabulary item 'die' has phoneme cluster /-fwa/ in Lenje and Soli. As mentioned above, the morphemes *-fwa, -mbi, -nwe, -nya* are equivalent morphemes in the languages given as examples.

Blount and Curley (1970) made no reference to sound correspondences. They considered those items where there was only one phonetic change, as cognates. For example, if a pair of items with an equal number of phonemes differed in one phoneme only, and if for example the difference was only in the feature 'voice,' that pair was considered to be

cognate. This study has borrowed Blount and Curley's criteria of cognates. Therefore, using Blount and Curley's (1970) criteria,

(a) Sala, *buzuba* and Soli, *busuba* both meaning 'day,' qualify as cognates despite having one phoneme which differs in voicing. The phoneme [z] from *buzuba* is a voiced alveolar fricative while the phoneme [s] from *busuba* is a voiceless alveolar fricative.

(b) The Lenje *ku-lya* and the Nyanja *ku-dya* 'to eat' qualify as cognates as they only differ in one phoneme. The phoneme [l] from *ku-lya* is an alveolar lateral approximant whereas the phoneme [d] from *ku-dya* is an alveolar plosive. Both [l] and [d] are voiced.

However, sometimes Blount and Curley's criteria seem rather extreme as they would disqualify pairs of items that otherwise would be mutually intelligible. For example,

(c) *Ku-bona* 'to see' in Ila, Lenje, Soli and Tonga; *ku-mona* in Kaonde; and *ku-ona* in Nyanja all qualify as cognates because of having one phoneme difference. The phoneme [β] is a bilabial fricative while the phoneme [m] is a bilabial nasal. In Nyanja there seems to be a deletion of a phoneme /-b-/ between /-u-/ and /-o-/, but the item is still regarded as cognate in this study because the languages have the same stem for the concept. But following Blount and Curley's criteria, the Nyanja *kuona* and the Ila, Lenje, Soli and Tonga *kubona* and Kaonde *kumona* on the other would not be cognates.

(d) The Soli k*u-kweela* and Tonga *ku-kwela* 'to pull' are regarded as cognates in this study regardless of the difference in vowel length manifested in Soli. Again, there

is no difference in voice, which may not be considered cognate using Blaunt and Curley's other criteria.

Nurse and Philippson (1980:26) took sound correspondences into account and considered two or more items to be cognates if they were derived from the same single item in a hypothetical common ancestor language by direct oral transmission. Borrowing Nurse and Philippson's criteria, the study shows that, for example,

(a) the Proto Bantu *–yàmà* 'animal' is *–nyama* and is cognate in all the eight languages being explored,

(b) the Proto Bantu numeral –na, -ne 'four' is cognate to Ila *syo-ne; Kaonde* kii-na; Lenje *sho-ne*; Nyanja *na-yi*; Sala *zho-ne*; Soli *chi-na-i* and Tonga *to-ne.*

(c) the Proto Bantu verb *–cek* 'laugh*'* is cognate to *ku-sek-a* in Ila*; ku-seek-a* in Kaonde*; ku-sek-a* in Lenje, Nyanja, Sala, Soli and Tonga*.* Although there is vowel lengthening in Kaonde *–seek-*, it is still regarded as cognate.

The lexical items in the above examples (a), (b), and (c) give evidence of sound correspondence, and are regarded as cognates by being derived from single items in Proto Bantu and taking sound correspondence into account.

Crowley (1992) based his criterion to determine cognates on vocabulary correspondence. He claimed that languages had lexical relationships because of the fact that they shared a Proto-Bantu form. For example, *mwana* /mu-ana/ 'child' in Nyanja, Sala, Soli and Tonga, which qualify as cognate because they all share the Proto Bantu form *-ana.*

Ethret (1996) considered shared innovations to determine cognates. He used morphemic and phonological features for the different languages that he classified. Further, he used the

stems or roots for the lexical items in order to know which lexical items originated from the same root and such words were regarded as cognates. For example, the Proto Bantu stem –yico 'eye' is found in the following languages: Ila: *liso*; Kaonde *jinso*; Lenje and Soli: liinso; Nyanja: *diso*; Sala: *linso*; Tonga: *liso*. Regardless of the different morphemic and phonological features of the item 'eye' in the languages studied, it is regarded as cognate. In Lenje, Sala and Soli, the stem –*iso* is nasalized to –*inso* while the rest of the languages have maintained the stem –*iso* although the initial phoneme is different in some languages. For example, the phoneme /y/ in Proto Bantu *yico* has changed to /l/ in Ila, Lenje, Sala, Soli and Tonga while it has changed to /d/ in Nyanja while the phoneme / tʃ/ from Proto Bantu -*yico* has changed to phoneme /s/ in all the seven languages investigated.

Another example, is the Proto Bantu root –*pa-* 'give' which is found in the following languages: Ila: *ku-pa*; Kaonde: *ku-pana*; Lenje: *ku-pa*; Nyanja: *ku-patsa*; Sala: *ku-pa*; Soli: *ku-pa*; Tonga: *ku-pa*. The lexical item may have different affixes, such as in Tonga, *ko-pa* 'give' , *nda-pa* 'I have given', *ba-pa* 'they have given', *ba-ba-pa* 'they have been given, or 'they have given them,' *wa-ndi-pa* 'he/she has given me', *ba-tu-pa* 'we have been given', *twa-pa* (*tu-a-pa*)'we have given,' denoting either tense or plural and singular, but as long as they have the same root, -*pa,* they would be regarded as cognates using Ethret's criteria.

Miti, L. M. (1996) showed that one of the problems in lexicostatistics is that of telling which roots are cognates and which ones are not in any pair of language varieties. He further demonstrated that there is no standard criterion to be used in judging cognation. Therefore, he proposed that the criteria for judging cognation should be relatively more inclusive so as to treat as cognates all those items that can correctly be interpreted by speakers of related languages without prior exposure. He further proposes that the

commonly accepted method of counting cognates in Swadesh lists should be based on the presumption of the genetic relation of two sets of lexemes. Further, he suggests that the relationship between two languages can be shown by demonstrating that two pairs are related to each other, and then counting the number of the cognate items in a vocabulary list.

Examples *d, e,* and *f* above suit Miti's (1996) criteria treated as cognate's items such as Swati *bita* and zulu *biza* 'call.' According to Miti (1995), although there is only one phoneme which differs, the difference is not only in voicing. The phoneme [t] is voiceless whereas the phoneme [z] is voiced, but the phoneme [t] is a plosive whereas the phoneme [z] is a fricative.

This study has employed Miti's proposed criteria for judging cognates, which is relatively more inclusive. It includes all the items that can be correctly interpreted or understood by speakers of other related languages. For example, *ku-belenga* 'to count' in Kaonde, Lenje, and *ku-werenga* 'to count' in Nyanja qualify to be cognates following Miti's inclusive criteria. The voiceless bilabial fricative [β] is represented by [w]in Nyanja, while the [l] is represented by [r] in Nyanja. Regardless of the above mentioned differences, the speakers of these languages can correctly understand the vocabulary items without prior exposure.

Another example is the item in Ila, Lenje, Sala and Soli: *ku-nwa* 'to drink.' The same item is represented as *ku-toma* in Kaonde, *ku-mwa* in Nyanja, *ku-nywa* in Tonga. The alveolar nasal [n] in Ila, Lenje, Sala, and Soli is manifested as the bilabial nasal [m] in Kaonde. The bilabial nasal [m] plus a labial velar approximant [w] making a cluster of [mw] in Nyanja is manifested as a three phoneme cluster [nyw] in Tonga. This cluster comprises the

alveolar nasal [ɲ], palatal approximant [j] and labial velar approximant [w]. Regardless of the above differences identified in different languages, it is still considered as a cognate in all the eight languages using Miti's inclusive criteria.

Table 3 below, shows more examples of the cognates identified from the seven languages under study following Miti's criteria for judging cognates.

Table 3: Cognate items based on Miti's (1996) criteria

English	Ila	Kaonde	Lenje	Nyanja	Sala	Soli	Tonga
Bad	ci-bi	-	ci-bi	-	ci-byabi	-	ci-bi
Bathe(v)	-	ko-vwa	-	-	-	-	-
Belly	-	ji-fumo	li-fumo	-	-vumo	-	-
Bird	-zuni	-ngonyi	-yuni	-	-yuni	-kakeeni	-yuni
Bite(v)	ku-luma	-	ku-luma	ku-luma	ku-luma	ku-luma	ku-luma
Blind person	moofu	Mpofu	Moofu	-	Moofu	Mpofu	Moofu
Bone	-	ni-kupa	-	-	-	chi-fupa	-
Breast	-	ji-bele	-	-bere	-	li-bele	-
Buy	ku-ula	-	ku-ula	ku-gula	ku-ula	ku-ula	ku-ula
Call(v)	-	ku-ita	ku-ita	ku-itana	ku-ita	-	ku-ita
Catch(v)	ku-kwata	ku-kwata	ku-cata	-	ku-jata	kwi-kata	ku-jata
Chew(v)	-	ku-nyeuna	ku-tafuna	ku-tafuna	ku-tafuna	ku-tafuna	ku-tafuna

Cloud(n)	-kkumbi	ji-kumbi	-nkumbi	-	-	-kumbi	-kkumbi
Come	-hhiza	-	-kosa	-	-konza	-kwesa	-
Day	bu-zuba	-juba	-	-	bu-zuba	bu-suba	bu-zuba
Sun(n)	i-zuba	-juba	li-suba	-dzuwa	i-zuba	li-suba	-zuba

For more details on cognates, see appendices A and B.

The items in the table above show the cognates that were judged according to Miti's criteria. For example the concepts for the word 'sun', *izuba* in Ila and Sala, *juba* in both Kaonde *lisuba* in both Lenje and Soli, *dzuwa* in Nyanja and *zuba* in Tonga, are regarded as cognates in all the seven languages. The speakers of these languages have no difficulties in understanding the concepts from other languages. However, there are notable differences that have been established. The [z] in Ila, Sala and Tonga is a voiced alveolar fricative, while the [s] in Soli and Lenje is a voiceless alveolar fricative. These sounds are both alveolar. The [j] in Kaonde and the [dʒ] in Nyanja are voiced post-alveolar affricates. However, the above differences do not affect the choice of an item as cognate in the languages explained above, using Miti's criteria.

Another example is the word 'catch' *kukwata* in Ila; Kaonde, k*ucata* in Lenje; *kujata* in Sala and Tonga; k*wikata* in Soli. However, despite the phonemic differences manifested in the forms in these languages we can still regard the forms as cognates, as far as Miti's criteria of selecting cognates is concerned.

4.3   Establishing Percentages of Inter-relatedness

The percentages of inter-relatedness between the languages studied were established from the 200 translated basic vocabulary items (Swadesh List). The basic vocabulary items which are listed in Appendix 'A' were analyzed and established cognates were coded. For cognates, a plus (+) sign was used while a minus (-) sign was used for items not judged to be cognates. A hash (#) sign was used for the second set of cognates while a caret (^) sign was used for a third set of cognates where in a list of seven items three groups of languages were found to be cognates. In order to determine the number of cognates in all the eight languages under study, counting of cognates was carried out after coding. The total number of cognates for each language was divided by two hundred (200) and then multiplied by one hundred (100) to establish the percentage of relatedness between the languages considered in this study. The data has been exemplified using   charts and tables. For more details on codification of the cognates see Appendix 'B'.

4.4   Results

In reference to the objectives outlined in 1.4 page page 19 of this study, the researcher used Miti's proposed criteria for judging cognates to established the percentages of the degree of inter-relatedness between the seven languages namely Ila, Lenje, Kaonde, Nyanja, Sala, Soli, and Tonga. The first objective was '**to establish the degree of inter-relatedness between Sala and the other Bantu Botatwe languages namely, Ila, Lenje, Soli and Tonga.**' The results/findings drawn from the analysis in Appendices 'A' and 'B' are presented and discussed below, and then the percentages of inter-relatedness were collapsed into a chart for Sala's inter-relatedness with other six languages studied inclusive of the Bantu Botatwe languages indicated in the above objective.

4.4.1. Sala

The degree of inter-relatedness between Sala and other languages under study was

analyzed as shown below.

Sala vs. Tonga   80%

Sala vs. Soli   67%

Sala vs. Lenje   81.5%

Sala vs. Ila   70%

Sala vs. Kikaonde   39.5%

Sala vs. Nyanja   39.5%

The above data is collapsed in the following Figure 5:
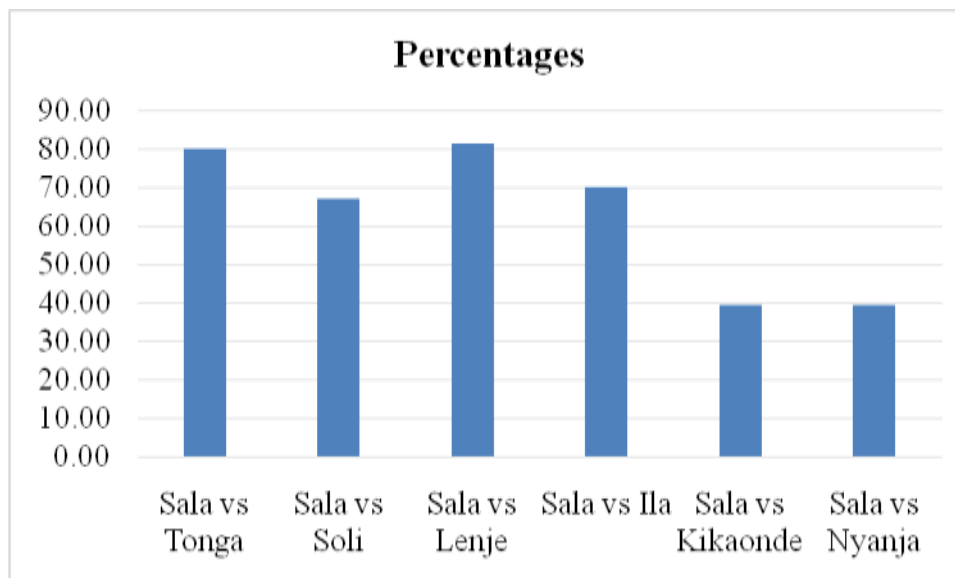
Figure 5: Sala's degree of inter-relatedness with other languages. (Source: Field work)

Figure 5 above is showing in graph form, the percentages of the degree of inter-relatedness between Sala and the other seven languages indicated in the graph. The highest percentage of inter-relatedness is between Sala and Lenje at 81.5 percent, followed by Sala and Tonga at 80 percent while Sala and Ila are at 70 percent. The lowest percentage in the above analysis is between Sala and Kikaonde, which is at 39.5 percent as well as Sala and Nyanja at 39.5 percent.

From the above data, it is evident that the degree of inter-relatedness between Sala and other Bantu Botatwe languages under study, particularly Tonga, Ila, Lenje and Soli is high, as it is above 67 percent.

As was pointed out in the background to this study, Tonga is taught to Sala children in schools. The figure above has shown that the degree of inter-relatedness between Sala and Tonga is at 80 percent, which is relatively high. Therefore, Sala children may have no major difficulties in learning Tonga since the degree of inter-relatedness between these two languages is high. It is a familiar language.

Kikaonde's degree of inter-relatedness with Sala is slightly below 40 percent. Ila, being part of Bantu Botatwe Group has a higher percentage of inter-relatedness with Sala than Kaonde. The study has shown that Sala fits into the concept of Bantu Botatwe due to its high percentage of inter-relatedness between Sala and the following languages belonging to Bantu Botatwe group: Lenje 81.5 percent; Tonga 80 percent; Ila 70 percent and Soli 67 percent.

The data analysis has shown that among all the seven languages under study, Sala is more closely related to Lenje, having a degree of inter-relatedness of 81.5 percent. The lowest

figures of inter-relatedness are between Sala and Kaonde at 39.5 percent; and between Sala and Nyanja, also at 39.5 percent. The data also provides evidence that Kikaonde has very little inter-relatedness with Sala, despite the fact that they share a geographical boundary.

4.4.2. Soli

The second objective was **'to ascertain how closely related is Soli to Nyanja.'** Drawing from the analysis presented in Appendices A and B, the following are the percentages of inter-relatedness between Soli and Nyanja including the other languages examined in this study:

Soli vs. Nyanja	45%
Soli vs. Ila	55%
Soli vs. Lenje	72%
Soli vs. Kikaonde	36.5%
Soli vs. Sala	67%
Soli vs. Tonga	63%

This data is collapsed in the following Figure 6:
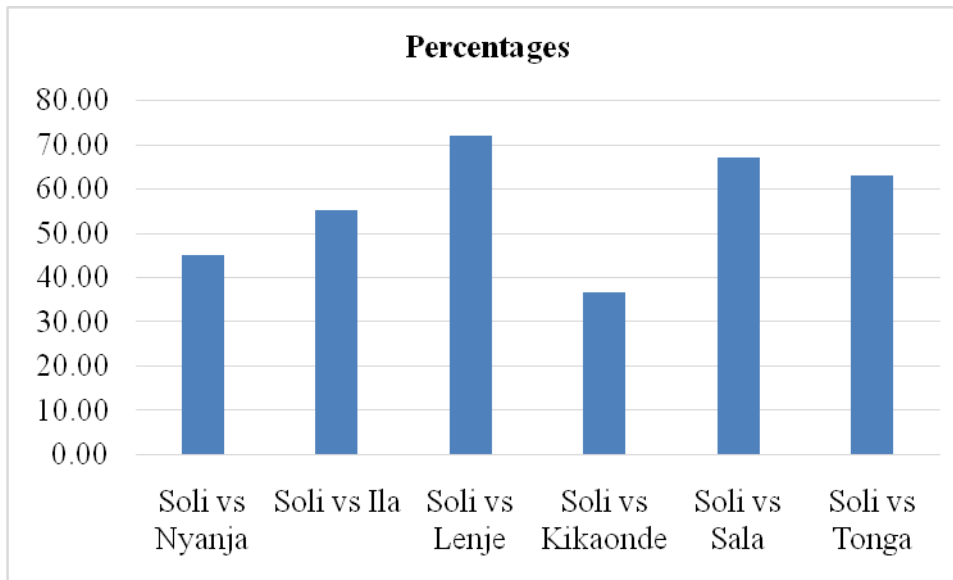
Figure 6: Soli's degree of inter-relatedness with other languages. (Source: Field work)

Figure 6 presents the degree of inter-relatedness between Soli and Nyanja, inclusive of the other languages investigated. While the degree of inter-relatedness between Soli and Nyanja is at 45 percent, the highest percentage of inter-relatedness in Figure 6 is between Soli and Lenje at 72 percent. It is followed by Soli and Sala at 67 percent. The third highest percentage of inter-relatedness is between Soli and Tonga at 63 percent, followed by Soli and Ila at 55 percent, The gap between the percentages of inter-relatedness for the Soli and Nyanja languages is too big to sustain the teaching of Nyanja as a school subject in Chongwe District. The Ministry of Education may wish to review the language policy following this revelation. The lowest percentage of the degree of inter-relatedness is between Soli and Kikaonde at 36.5.

4.4.3. Kikaonde

The third objective of this study was '**to ascertain how closely Kaonde is related to Tonga**.' The analysis in Appendices 'A' and 'B' revealed the following percentages of the

86

degree of inter-relatedness between Kikaonde and Tonga, including the other languages studied:

Kikaonde vs. Tonga   33%

Kikaonde vs. Ila   30%

Kikaonde vs. Sala   39.5%

Kikaonde vs. Nyanja   32%

Kikaonde vs. Lenje   41%

Kikaonde vs. Soli   36.5%

This data is collapsed in the following Figure 7:
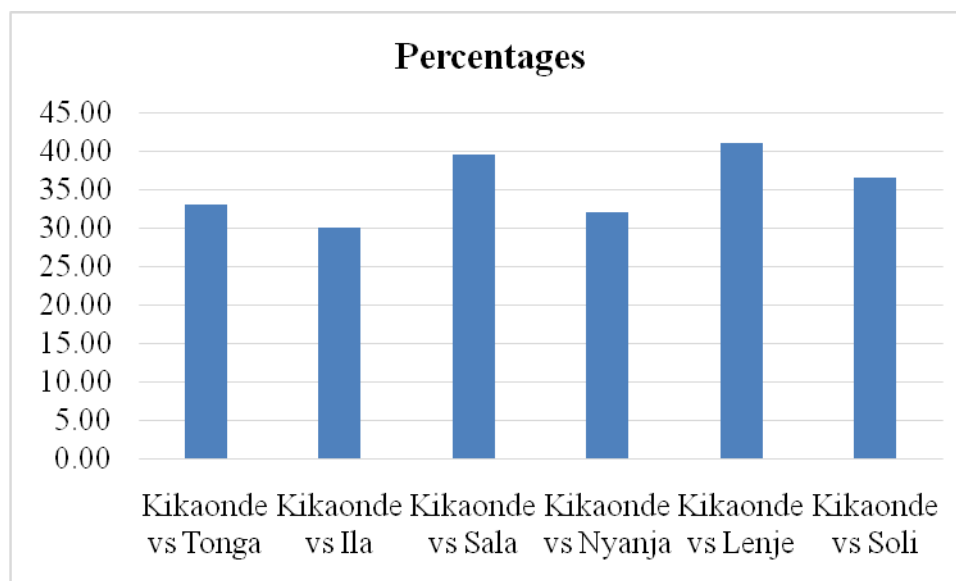
Figure 7: Kikaonde's degree of inter-relatedness with other languages.

(Source: Field work)

In Figure 7 the graph displays the degree of inter-relatedness involving Kikaonde and Tonga. The other languages investigated were also included. The data in the graph above is quite different from the data in the graphs that have been discussed above. While the percentage of the degree of inter-relatedness between Kikaonde and Tonga is at 33 percent, the highest percentage of inter-relatedness is between Kikaonde and Lenje at 41 percent. The second highest percentage of inter-relatedness is between Kikaonde and Sala at 39.5 percent, which is followed by Kikaonde and Soli at 36.5 percent. The second lowest percentage of inter-relatedness is between Kikaonde and Nyanja at 32 percent while the least percentage of inter-relatedness is between Kikaonde and Ila at 30 percent.

From the above analysis in Figure 7, the researcher observed that while the highest percentage of inter-relatedness is between Kikaonde and Lenje, Tonga is taught as a subject in the Kaonde Speaking area. The gap between the percentages of inter-relatedness for the two languages (Kikaonde and Tonga) is too big to support the teaching of Tonga as a school subject. During the oral interviews held on 19th February 2013, with His Royal Highness, Chief Mumba, at his palace, the chief confirmed that there is a lot of confusion in terms of language. He said that Kaondes, Ilas and Tongas use Tonga as a language of literacy and a school subject. The chief pleaded that the language planners may consider Kaonde as a school subject in his chiefdom. And this concern is supported by the study in that it has revealed that the percentage of inter-relatedness between Tonga and Kaonde is at 33 percent. Such a low percentage might affect the performance of the Kaonde pupils in schools while Ilas may not have any challenges because the percentage of inter-relatedness between Tonga and Ila is at 76.5 percent. Hence the Ministry of Education may wish to

instead introduce or use Kikaonde as the language of literacy and Zambian language subject in the Kaonde speaking area.

The study has also established that among the seven languages studied, Lenje and Tonga have the highest percentage of retention of Proto Bantu lexical items at 32 percent. The Proto Bantu lexical items retention may have contributed to the high percentage rate of inter-relatedness between Lenje and Tonga.

4.4.4   Tonga

The percentage of the degree of inter-relatedness between Tonga and the rest of the languages under study was analyzed and results are presented below.

Tonga vs. Ila   76.5%

Tonga vs. Lenje   84.5%

Tonga vs. Kikaonde   33%

Tonga vs. Sala   80%

Tonga vs. Soli   63%

Tonga vs. Nyanja   40.5%

The above data is illustrated in Figure 8 bellow:
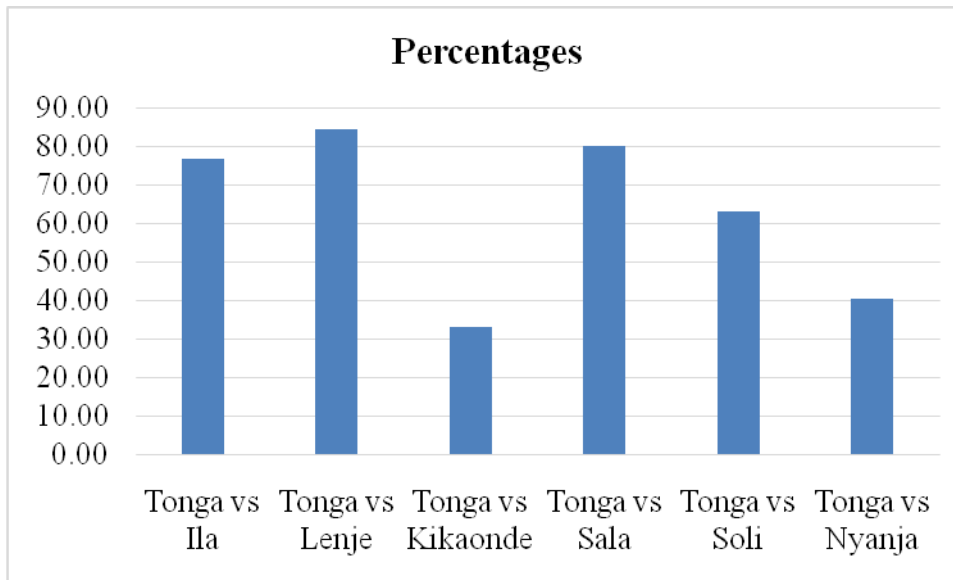
Figure 8: Tonga's degree of inter-relatedness with other languages.

(Source: Field work)

The above figure displays the degree of inter-relatedness between Tonga and the other languages investigated. The highest percentage of inter-relatedness in Figure 8 is between Tonga and Lenje at 84.5 percent followed by Tonga and Sala at 80 percent. The third highest percentage of inter-relatedness is between Tonga and Ila at 76.5 percent, followed by Tonga and Soli at 63 percent, while Tonga and Nyanja is at 40.5 percent. The second lowest percentage of inter-relatedness is between Tonga and Kaonde Ila at 37.5 percent. The least degree of inter-relatedness is between Tonga and Kikaonde at 33 percent. Tonga is the Zambian language taught in the Lenje speaking area. From the analysis shown above, the percentage of inter-relatedness between Tonga and Lenje is relatively high. This means that Lenje learners do not have challenges in learning Tonga, because most of the basic vocabulary is similar in the two languages. The researcher supports the language policy makers for choosing Tonga as a school subject in the Lenje speaking area. It is a familiar language.

4.4.5. Ila

The study has established the percentages of inter-relatedness between Ila and the other seven languages examined as follows:

Ila vs. Sala   70%

Ila vs. Tonga   76.5%

Ila vs. Lenje   66.5%

Ila vs. Soli   55%

Ila vs. Nyanja   30%

Ila vs. Kikaonde   30%

These data are illustrated in Figure 9:
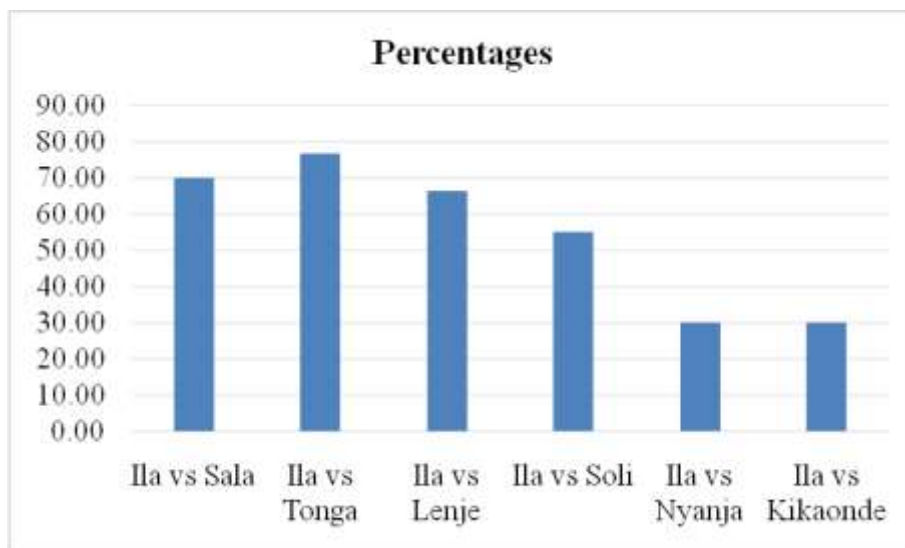
Figure 9: **Ila's degree of inter-relatedness with other languages.** (Source: Field work)

Figure 9 presents the degree of inter-relatedness between Ila and the other languages investigated. The highest percentage of inter-relatedness is between Ila and Tonga at 76.5 percent, followed by Ila and Sala at 70 percent. The third highest percentage of inter-relatedness is between Ila and Lenje at 66.5 percent, while that of Ila and Soli is at 55

percent. The lowest degree of inter-relatedness is between Ila and Kaonde at 30 percent, as well as Ila and Nyanja also at 30 percent.

4.4.6. Lenje

The following are the percentages of inter-relatedness between Lenje and the other languages investigated in this study:

Lenje vs. Tonga   84.5%

Lenje vs. Soli   72%

Lenje vs. Sala   81.5%

Lenje vs. Nyanja   40%

Lenje vs. Kikaonde   41%

Lenje vs. Ila   66.5%

This data is collapsed in the following Figure 10:
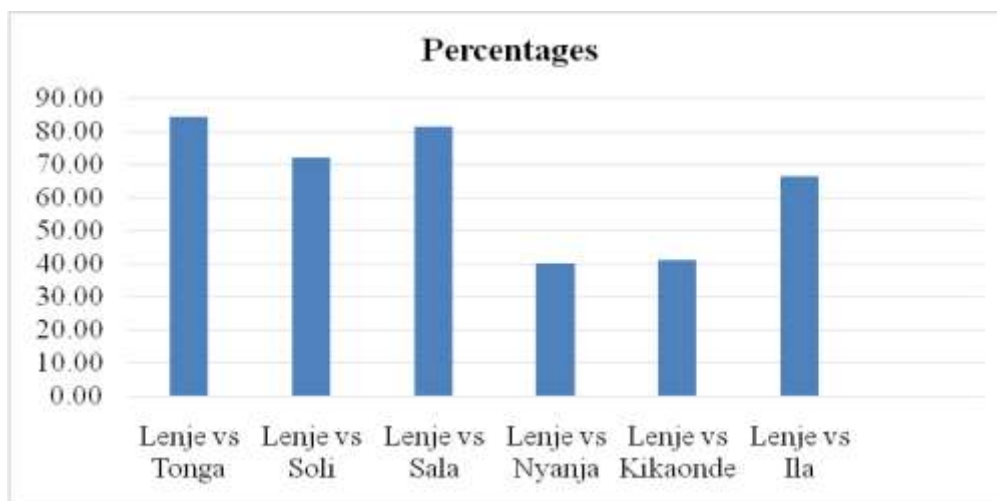
Figure 10: Lenje's degree of inter-relatedness with other languages. (Source: Field work)

In Figure 10 the graph presents the degree of inter-relatedness between Lenje and the other languages investigated. The highest percentage of relatedness in Figure 10 is between Lenje and Tonga at 84.5 percent. This is followed by Lenje and Sala at 81.5 percent. The third highest percentage of inter-relatedness is between Lenje and Soli at 72 percent, which is followed by Lenje and Ila at 66.5 percent while Lenje and Kaonde are at 41 percent. The sixth lowest percentage of inter-relatedness is between Lenje and Nyanja at 40 percent while the lowest percentage of inter-relatedness is between Lenje and Kaonde Ila at 39.5 percent.

4.4.7 Nyanja

The following are the percentages of the degree of inter-relatedness between Nyanja and other languages studied:

Nyanja vs. Sala   39.5%

Nyanja vs. Soli   45%

Nyanja vs. Ila   30%

Nyanja vs. Lenje   40%

Nyanja vs. Tonga   40.5%

Nyanja vs. Kaonde   32%

This data is collapsed in the Figure 11 below:
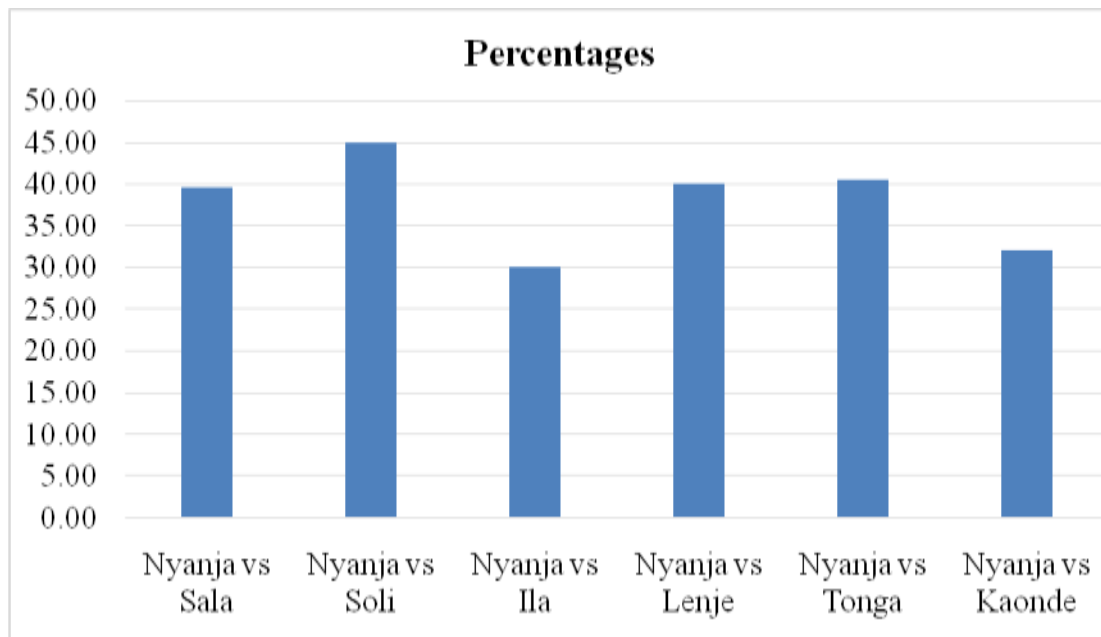
Figure 11: Nyanja's degree of inter-relatedness with other languages.

(Source: Field work)

In Figure 11, the data displays the degree of inter-relatedness involving Nyanja and the other languages investigated. The graph above is quite different from the graphs that have been discussed before, because all the percentages of inter-relatedness are below 50. The highest percentage of inter-relatedness is between Nyanja and Soli at 45 percent. This is followed by Nyanja and Tonga at 40.5 percent, followed by Nyanja and Lenje at 40 percent. The fourth highest percentage of inter-relatedness is between Nyanja and Sala at 39.5 percent, followed by Nyanja and Kikaonde at 32 percent. The lowest percentage of inter-relatedness is between Nyanja and Ila at 30 percent.

The data in Figure 11 shows that among all the seven languages under study, Nyanja has the lowest degree of inter-relatedness with the other languages, ranging from 30 and 45

percent. The highest degree of inter-relatedness in Figure 9 is between Nyanja and Soli at 45 percent and the lowest is between Nyanja and Ila, which is at 30 percent. Although Nyanja has the lowest degree of inter-relatedness to all the languages under study, it was rated number two in terms of predominant language of communication. According to the Census Report (2010:63) 14.8 percent of the population in the country spoke Nyanja as the language of communication. The report shows that Nyanja is spoken in all the nine provinces that existed in the year 2010. However, the highest percentage of speakers is found in Lusaka Province at 61.9 while the second highest is in Eastern Province at 17.4 percent. The third highest percentage of speakers of Nyanja is found in Central Province at 8.9. (Source: Census Report 2010:66.)

While Nyanja is used for literacy and a school subject in the Soli speaking area, the percentage of inter-relatedness between Nyanja and Soli is at 45 percent. This is a low percentage especially when one considers that Soli has a high percentage of inter-relatedness between other languages within the Bantu Botatwe group. For example, the inter-relatedness between Soli and Lenje is at 72 percent; Soli and Sala are at 67 percent while Soli and Tonga is at 63 percent. This kind of discrepancy would cause challenges to learners who use Nyanja in the Soli speaking area as languages of initial literacy due to the low percentage of inter-relatedness. However, the researcher examined the learners of Nyanja in Chongwe district to check their performance. The results are presented in chapter 5.

4.5   Group Average Method

On the basis of the percentages of the inter-relatedness of the languages under study, the group average method was employed to try and ascertain the degree of inter-relatedness

between these eight languages. The lexicostatistical percentages are presented in Table 4 below where the eight languages have been abbreviated due to space limitation as follows: SA representing Sala, IL representing Ila, KA representing Kaonde, LE representing Lenje, NY representing Nyanja, SO representing Soli and TG representing Tonga.

Table 4: Lexicostatistical percentages of inter-relatedness

TG	TG						
LE	**84.5%**	LE					
KA	33%	41%	KA				
IL	76.5%	66.5%	30%	34.5%	IL		
SA	80%	81.5%	39.5%	42%	70%	SA	
SO	63%	72%	36.5%	39.5%	55%	67%	SO	
NY	40.5%	40%	32%	31.5%	30%	39.5%	45%	NY

From the above percentages of the 200 common basic vocabulary items, it has been observed that Lenje and Tonga share the highest percentage of inter-relatedness, which is at 84.5 percent, while Ila and Kaonde; Ila and Nyanja share the lowest percentages, which is 30 percent for each pair. The general observation from the study is that, Nyanja is least related to any of the other languages under study. Indeed, it is not one of the Bantu Botatwe languages.

Using the Group Average Method, that group languages together into a table with their inter-relatedness, the two languages with the highest inter-relatedness are grouped into one column in the next table, where the percentage of inter-relatedness is established among

the two highest languages and the other languages being studied. Thereafter the language with the higest percentage is added to the first two to make three highest inter-related languages amongst all languages under investigation. The process goes on until the lowest in the group is established. The hierarchy of inter-relationship between the eight languages under study was worked out. As a result subsequent tables were made to present the averages of lexicostatistical percentages between the most closely related languages.

After collapsing Table 4 focussing on the columns for Lenje and Tonga which have the highest percentage of 84.5 percent (refer to Table 4) we obtained the follwong average percentages.

Table 5: Group Average Method 1

TG /LE	TG /LE						
KA	37%	KA					
IL	71.5%	30	34.5%	IL			
SA	**81%**	39.5%	42%	70	SA		
SO	67.5%	36.5%	39.5%	55%	67%	SO	
NY	40.2%	32%	31.5%	30%	39.5	45%	NY

As Table 5 above shows the highest percentage of inter-relatedness is now between Tonga, Lenje and Sala at 81 percent. The highest percentage is shown in bold in the above table. The next step is to collapse Table 5 by working out the group averages percentages

between the languages with the highest degree of inter-relatedness that is Tonga and Lenje on one hand and Sala on the other. These group averages are presented in Table 6 below.

Table 6: Group Average Method 2

TG/LE/SA	**TG /LE/SA**					
KA	60.3%	**KA**				
IL	**70.8%**	30%	34.5%	**IL**		
SO	67.3%	36.5%	39.5%	55%	**SO**	
NY	41.1%	32%	31.5%	30%	45%	**NY**

The highest percentage of inter-relatedness in Table 6 is between Tonga, Lenje, Sala and Ila at 70.8 percent, and is shown in bold in the above table. In Table 7 we provide the group averages collapsing the columns for Tonga, Lenje, Sala and Ila as follows:

Table 7: Group Average Method 3

TG /LE/SA	**TG/LE/SA/IL**			
KA	51%	**KA**		
SO	**67.3%**	36.5%	**SO**	
NY	41.1%	32%	45%	**NY**

As can be seen from Table 7 above the highest degree of inter-relatedness is now between Tonga/Lenje/Sala/Ila and Soli on the other hand at 67.3 percent. The group averaging of these languages gives us Table 8 below.

Table 8: Group Average Method 4

TG /LE/SA/IL	TG /LE/SA/IL/SO		
KA	**44.5%**	KA	
NY	43.1%	31.7%	NY

In Table 8, the highest percentage of inter-relatedness is between Tonga, Lenje, Sala, Ila, soli on one hand and Kaonde on one other at 44.5 percent. Further group averaging based on the languages with the highest degree of inter-relatedness that is Tonga/Lenje/Sala/Ila/Soli and Kaonde results in the percentages presented in Table 9 below.

Table 9: Group Average Method 5

TG /LE/SA/IL/SO/KA	TG /LE/SA/IL/SO/KA	
NY	38.1%	NY

In Table 9 the highest percentage is between TG/LE/SA/IL/SO and KA at 44.5 percent.

Tables 9 above shows that the least related languages in the category of languages under study are Kaonde and Nyanja. From Table 9 it can be seen that the group average percentage of inter-relatedness between TG/LE/SA/IL/SO/KA and NY is at 38.1 percent. Table 9 also shows the language that is least related to the others in the group of languages we are investigating which is Nyanja. Nyanja is only related with a percentage of 38.1percent to all the languages we have investigated, that is Tonga, Lenje, Sala, Ila, Soli,

Kaonde. The hierarchy of the inter-relatedness of the languages based on the group averaging above can be represented as in the tree diagram in Figure 13:
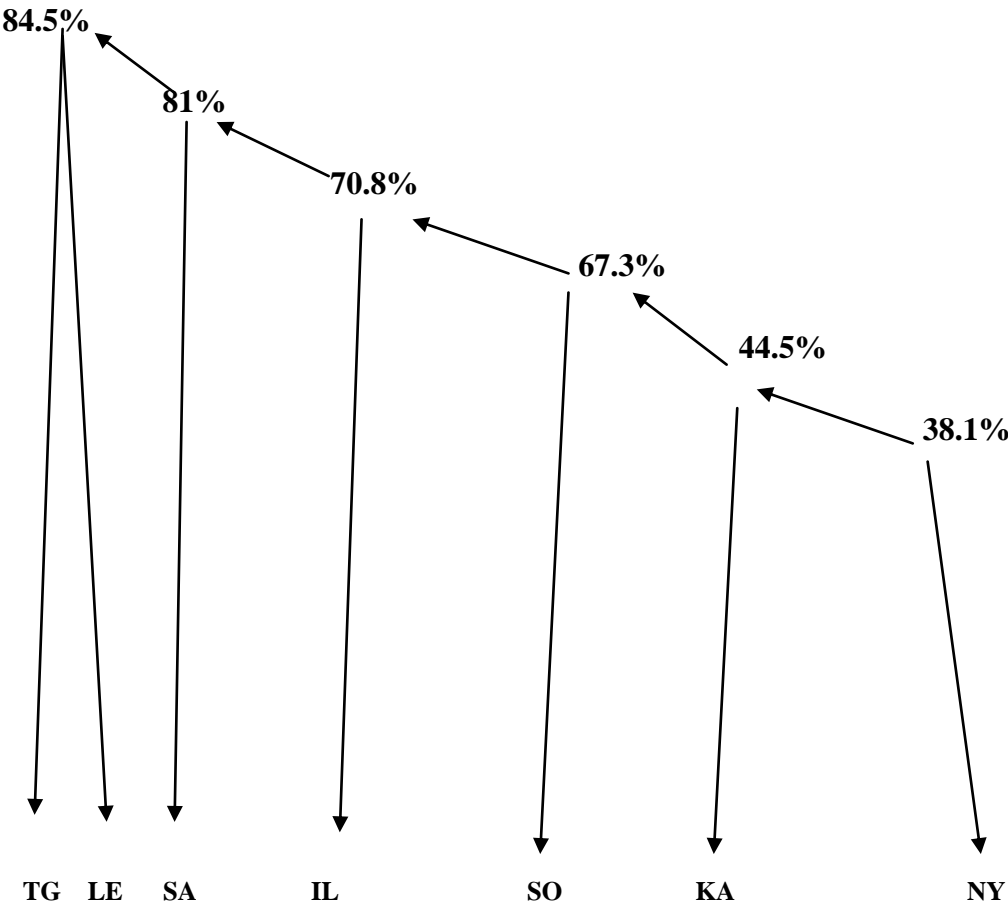
Figure 12: Hierarchical Group Average Percentages of Inter-relatedness

Figure 12 is a tree diagram illustrating the hierarchical group average inter-relatedness of Tonga, Lenje, Sala, Ila, Soli Kaonde, Kaonde Ila and Nyanja. Lenje and Tonga are the most highly related among the languages studied with 84.5 percent, seconded by Sala at 81 percent. The group average percentage of inter-relatedness between Tonga, Lenje, Sala and Ila is at 70.8 percent, followed by Soli at 67.3percent.The group average for KA and the rest of the languages investigated is at 44.5 percent while the group average percentage between Nyanja and the other languages investigated is at 38.1percent. From the analysis

shown on Figure 12, there is a possibility that the two highest inter-related languages are the dialects of Tonga.

On the other hand, Soli shares a high percentage of inter-relatedness with Lenje, at 72 percent. The researcher observed that the high levels of inter-relatedness between these two languages are characterized by the close proximity of the speakers because they share boundaries between the Great East road and the Great North Road. The preliminary investigations at the Curriculum Development Center showed that some years ago, Lenje was used for literacy and as a school subject in Soli speaking area.

4.6.   Cognate Retention

The fourth objective of this study was **'to ascertain whether there is any retention of the Proto-Bantu vocabulary in the languages under investigation.'**

In this study, cognate retention concerns cognates that are found in Ila, Kaonde, Lenje, Sala, Nyanja, Soli, and Tonga, that emanated from Proto Bantu. The data in Appendix 'C' showing cognates retained from Proto Bantu was analyzed and the group of lexical items that was close to Proto Bantu was coded with a plus (+) sign while the other groups of cognates that were different from Proto Bantu were coded with a hash (#) sign. After coding the data, analysis was done in order to establish the percentages of cognate retention from Proto Bantu in all the seven languages studied.

In reference to data analysis presented in Appendix 'C', the study has shown that numerals are part of the retained items from Proto Bantu. For example, the numerals, one, two, three, four, five, and ten are cognates in all the languages under study and have been retained from Proto Bantu.

Table 11below shows the above mentioned numerals:

Table 10: Cognates retained from Proto Bantu for numerals

Gloss	One	Two	Three	Four	Five	Ten
Proto Bantu	-mó	-bidi	-catu, -tatu	-na, -ne	-caanu, -taanu	-kume, -kumi
Ila	-mwi	-bile	-tatwe	-ne	-sanwe	-kumi
Kaonde	-mo	-biji	-satu	-na	-tanu	-kumi
Kaonde Ila	-mo	-biji	-satu	-na	-tanu	-kumi
Lenje	-mwi	-bilo	-tatwe	-ne	-sanu	-kumi
Nyanja	-mo-	-wiri	-tatu	-na	-sanu	-khumi
Sala	-mwi	-bilo	-tatwe	-ne	-sanu	-kumi
Soli	-mo	-bili	-tatu	-na	-sanu	-kumi
Tonga	-mwe	-bilo	-tatwe	-ne	-sanwe	-kumi

It should be noted that all examples are presented as stems. The cognates for numerals presented in Table 10 have various phonological forms. The following are examples:

a) The Proto Bantu –mó [mo] 'one' is –mwi [mwi] in Ila and Sala while it is –mo [mo] in Kaonde, Nyanja and Soli. In Tonga, it is –mwe [mwe].

b) The cognate for Proto Bantu –bidi [βidi] 'two' is –bili [βili] in Soli; -bile [βile] in Ila; -bilo [βilo] in Lenje , Sala and Tonga; -biji [βidʒi] in Kaonde; -wiri [wiri] in Nyaja.

(c) The item for Proto Bantu –catu [tʃatu], -tatu [tatu] 'three' manifests phonological differences in the languages under study. In Ila, Lenje, Sala and Tonga, it is tatwe [tatwe] while in Kaonde it is –satu [satu]. In Nyanja and soli it is –tatu [tatu].

(d) The Proto Bantu item –na, -ne [na , ne] 'four' is presented in Kaonde, Nyanja, and Soli, as –na [na] while in Ila, Lenje, Sala and Tonga, it is –ne [ne].

(e) The Proto Bantu –caanu [tʃaanu], taanu[taanu], 'five' is reflected as –sanwe [sanwe] in Ila and Tonga while in Lenje, Nyanja, Sala and Soli, it is –sanu [sanu] and –tanu [tanu] in Kaonde.

(f) The Proto Bantu item –kume [kume], -kumi [kumi],is reflected as –kumi [kumi] in Ila, Kaonde, Lenje, Sala, Soli and Tonga; while in Nyanja it is –khumi [kʰumi].

The researcher observed that in example (a) the following phonological differences have been realized: in Ila and Sala the phoneme glide /w/ has been inserted after /m/ to create the cluster /mw/, the vowel /o/ has been raised to /i/, thus –*mo* →-*mwi.* In Tonga the vowel /o/ has been fronted to /e/. Thus –mo →-*mwe*.

In example (b) the Proto Bantu phoneme /d/ from [βidi] has been literalized to /l/ in Ila, Lenje, Sala, Soli and Tonga. The vowel /i/ after the phoneme /d/ has been changed as follows:

Language	Item	Change
Soli	bile [βile]	/i/ → /e/ (involving vowel lowering)
Ila,Lenje, Sala, and Tonga	bilo [βilo]	/i/ → /o/ (involving vowel lowering)
Kaonde	biji [βidʒi]	/i/ has been maintained

In Kaonde the Proto Bantu phoneme /d/ has been affricativised to/ʤ/.

In example (f) there is a difference from the Proto Bantu item –kumi which involves aspiration in Nyanja –khumi [kʰumi]. These are the few examples of phonological differences established in the study. In (c) / ʧ/ → t in Ila, Lenje, Sala, Tonga and Nyanja. Similarly, / ʧ/ → t in Kaonde. Another notable change affects /u/ which is refleced as /e/. Further there the insertion of /w/ before /e/ in Ila, Lenje, Sala, Tonga and Nyanja.

In (e) /ʧ/ → /s/ in Ila, Tonga, Lenje, Nyanja, Soli and Sala. There is also vowel shortening in all the languages concerned. For Ila and Tonga there is further /w/ insertion.

Names of animals, such as dog, elephant, are also cognates retained from Proto Bantu in all the languages with the exception of Nyanja, which has a non-cognate. It should be noted that the two names of animals are the only ones included in the sample data in cognate retention. Therefore, only two names of animals have been used in table 12 below showing the retained cognates.

Table 11: Retained Animal Names from Proto Bantu

Gloss	Dog	Elephant
Proto Bantu	-bóà	-jogu
Ila	-bwa	-zohhu
Kaonde	-bwa	-nzovu
Lenje	-bwa	-nsofu
Nyanja	Galu	-njovu
Sala	-bwa	-zovu
Soli	-bwa	-jofu
Tonga	-bwa	-zovwu

The data presented in Table 11 above shows that the Proto-Bantu item *–boa* has not changed. The combination of the vowels /o/ and /a/ as *oa* is phonetically pronounced as [w] in some Bantu languages such as Sesotho where words such as *bajoa* 'get frosted' (plants), *joalo* 'like that', *koahela* 'to cover' are transcribed as [βadʒwa], [dʒwalo] and [kwahela] respectively. Hence, the researcher is assuming that the Proto Bantu form βoa is phonetically [βwa].The Proto Bantu item–*njogu* [ndʒogu] 'elephant' has undergone change through phonological processes although it has been retained as a cognate in all the languages under study. In Kaonde, Kaonde-Ila and Sala, an 'elephant' is *-nzovu*; in Lenje it is called *-nsofu*. In Nyanja it is called *–njovu* while in soli it is *–jofu*. The Proto-Bantu item *jogu* 'elephant' has undergone the following phonological change: in Nyanja *njovu* the voiced velar plosive /g/, has been fricativized to the phoneme /v/which is a voiced labio-dental-fricative. In addition, the Proto Bantu [dʒ] has been pre-nasalized to [ndʒ]. In Soli, the voiced Proto Bantu plosive /g/ is fricativised to a voiceless labio-dental-fricative /f/. In Lenje, the Proto Bantu affricate [dʒ] has been fricativised to the voiceless alveolar fricative phoneme/s/ and then pre-nasalized to /ns/. Also, the voiced velar plosive /g/ has been fricativised to a voiceless labio-dental fricative /f/. Other changes include the following:

/j/ → /z/ (Tonga, Sala, Ila, Kaonde,)

Prenasalization of /z/ → /nz/ in Kaonde.

/g/ → /v/ in Kaonde, Sala and Tonga

/w/ insertion in Tonga

/g/ → /f/ in Soli

/g/ → /ɦ/ in Ila

However, regardless of the above phonological changes, the Proto Bantu lexical item *jogu* is considered retained in all the seven languages studied because the item is intelligible to the speakers of these eight languages following Miti's (1996) criteria of cognates.

Concepts for the body parts such as ear, eye, head, thigh and tooth have also been retained from Proto Bantu in all the languages under study apart from one item 'thigh' which is non-cognate in Kaonde, Kaonde Ila and Nyanja.

Table 12 below displays the cognates of body parts as follows:

Table 12: Cognates of Body Parts retained from Proto Bantu

English	Ear	Eye	Head	Thigh	Tooth
Proto Bantu	-to	yi-co	-tó	-bÈdÒ	yí-Nò
Ila	-twi	li-so	-twi	-belo	li-no
Kaonde	-twi	ji-nso	-twe	-joma	ji-no
Kaonde Ila	-twi	ji-nso	-twe	-jooma	ji-no
Lenje	-twi	li-inso	-twi	-belo	li-ino
Nyanja	-tu	di-so	-tu	-ncafu	dzi-no
Sala	-twi	li-nso	-twi	-belo	li-no
Soli	-twi	li-inso	-twi	-belo	li-no
Tonga	-twi	li-so	-twe	-belo	li-no

Table 12 shows that while the body parts included in the 200 basic vocabulary list have been considered as cognates, there are some phonological differences in the lexemes noted

106

between the ancestor language and the languages invested. For example, the Proto Bantu *-to* 'ear' is *–tu* in Nyanja which shows a phonological change that has taken place between the Proto Bantu and Nyanja. The vowel /o/ has been raised to /u/while the phoneme /t/ is retained. In the other seven languages it is *–twi*, showing the insertion of /w/ after /t/. Then the vowel /u/ has been fronted to vowel /i/ which results in gliding to /tui/ '*twi*.'

The Proto Bantu *–yico* 'eye' is cognate with *–liso* in Ila and Tonga although it has experienced a phonological change as well. The Proto Bantu phoneme /j/ has lateralized to /l/ while the phoneme /ʧ/ has changed to the fricative /s/.The Proto Bantu item *–yico* 'eye' is also cognate with *–jiso* in Kaonde and Kaonde Ila despite the phonological change where /y/ which has been affricativized to [ʤ] while the affricate /ʧ/ has changed to the fricative /s/. The observation is that /ʧ/ has been fricativised to /s/ in all the languages investigated.   In Lenje the item has a long vowel/i: / in*–liinso* while in Nyanja the gliding /l/ is realized as the plosive /d/, hence the item becomes *–diso*.

For the Proto Bantu item *–to* 'head' the cognate in Ila, Lenje, Sala, and Soli is *–twi*. In Kaonde, Kaonde Ila and Tonga it is *–twe.* Similarly, the Proto Bantu item-*to* 'head' is *–tu* in Nyanja suggesting that a phonological change has taken place between the Proto Bantu and Nyanja. The mid back rounded vowel /o/ has been changed to the high back rounded vowel /u/while the phoneme /t/ is retained.   In the other seven languages it is *–twe*, showing a phonological change of the insertion of /w/ after the phoneme /t/ to form the cluster /tw/. Then the mid back rounded vowel /o/ has been changed to the mid front unrounded vowel /e/.

The foregoing examples and the rest of items in Table 12 have been considered cognates using Miti's (1996) inclusive criteria for identifying cognates. This means that the items are intelligible and the speakers of these languages are able to understand the items without prior knowledge of the language.

Appendix C shows more items that have been retained as cognates from Proto Bantu. However, the study has also shown that there are some cognate items within the languages under study that are not derived from Proto Bantu.For example, the concept for 'smoke' is cognate in all the languages investigated but is not derived from Proto Bantu. Also, the concept for 'hear' is cognate in all the eight languages investigated but does not originate from Proto Bantu.

It has been observed from the data analysis that out of the 200 vocabulary items in the word list used for the study, 101 are Proto Bantu forms. It has also been established that 32 items out of the 101 Proto Bantu forms, are cognates in all the languages under study. Out of the 32 cognates, 25 vocabulary items have been retained from Proto Bantu. However, out of 101 Proto Bantu Forms there are some items which are cognates in two, three, or four languages while the same items are non-cognates in other languages under study. The following table displays the Proto Bantu vocabulary retention percentages of each language:

Table 13: Retention Percentage of Proto Bantu forms

Language	Lexical items retained	Retention Percentage
Ila	57	28.5%
Kaonde	57	28.5%
Lenje	64	32%
Nyanja	56	28%
Sala	60	30%
Soli	54	27%
Tonga	64	32%

Table 13 has reveals that Tonga and Lenje have the highest percentage of retention of Proto- Bantu forms at 32 percent. The second highest is Kaonde Ila at a 30.5 percent retention rate, while Sala is in the third position with at 30 percent retention rate.Kaonde and Ila are in the fourth position with a 28.5 retention percentage. Nyanja is in the fifth position with 28 percent of retention rate. The sixth and last is Soli with a percentage of 27.

The percentages retained from Proto Bantu Forms are shown in Figure14 below:
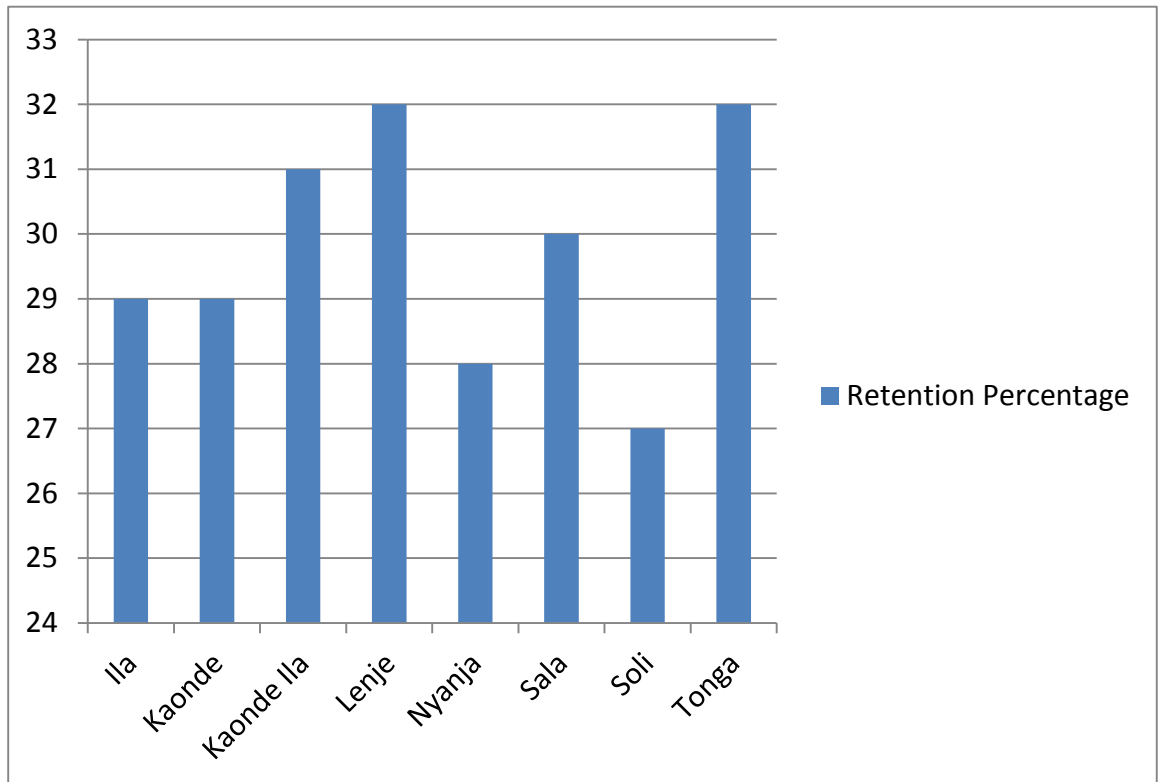
Figure 13: Retention Percentage of the Proto Bantu (Source: Field Work**)**

Figure 13 displays the levels of the percentages of the retention of Proto Bantu forms by the eight languages investigated.

4.7   Shared Cognates

Shared cognates refer to items that are cognates but they do not originate from Proto Bantu.

Appendix C presents cognates that do not appear in all the languages as cognates, but appear in groups of two, three or four as shared cognates, suggesting that some shared cognates do not originate from Proto Bantu while some do.   Hillary (2009) cites Hinnesbusch (1999:177), that

lexicostatistics has a potential for providing evidence of contact. He explains that the operative assumption here is that similarities which are measured by lexicostatistics are more than an indication of genetic transmission, and that such similarities are of multiple origins. He asserts that it has been recognized that vocabulary is the linguistic component which travels most easily from dialect to dialect, from one related adjacent language to another, and from genetically distant or unrelated adjacent language to another... vocabulary is the linguistic component which travels most easily from dialect to dialect.

Following Hinnesbusch's line of thought stated above, the study has shown that some languages are related due to theirhistorical background, not necessarily originating from Proto Bantu. For example, it is mentioned in the background to the study that Sala people may have originated from Gwembe, meaning that historically they were Tonga people. This study has established that the percentage of degree of relatedness between Sala and Tonga is 80 percent. Most of the lexical items that are cognates between these two languages do not originate from Proto Bantu. This study has shown that the languages explored, particularly those under Bantu Botatwe group have above 50 percent lexical agreement due to language contact.

The following table shows evidence of the lexical items that are related but have no cognates with Proto Bantu:

Table 14: Cognates not Retained from Proto Bantu

Gloss	Bad	Smoke	Hear	Louse	Pull
Proto Bantu	-béép-	-yoki	-yigu	-dá	Dìòd
Ila	-bi	bushi	-mvwa	njina	-kwela
Kaonde		bwishi	-mvwa		
Kaonde Ila		bwishi	-mvwa		
Lenje	-bi	bwishi	-mfwa	njina	-kwela
Nyanja		Utsi	-mva	-nda	
Sala	-bi	bwishi	-mva	njina	
Soli		bwishi	-mfwa	njina	-kweela
Tonga	-bi	Busi	-mvwa	njina	-kwela

Table14 shows lexical items that are shared between two, three or more languages but
which do not originate from Proto Bantu. For example, the word 'louse' is *njina* in Ila,
Lenje, Sala, Soli and Tonga, yet in Proto Bantu it is *–dá.* In Nyanja it is *–nda*, which
derives from Proto Bantu and /d/ is prenasalized. The lexical item for 'pull' in Proto Bantu
is *Dìòd* while in Ila, Lenje, Soli and Tonga, it is *–kwela* although vowel lengthening
applies in Soli to realize *-kweela*. Some lexical items like 'hear' are similar in all the
languages under study save for a few slight variations but do not originate from Proto
Bantu. The Proto Bantu item *–yigu* 'hear' is *–mvwa* in Kaonde, Kaonde Ila and Tonga,
while In Lenje and Soli it is *–mfwa*. In Nyanja and Sala it is mva. Although there are some
phonological differences in the lexeme among the eight languages such as the voiceless
labio-dental fricative /f/ in Soli, and Lenje, while the other languages have a voiced labio-

dental fricative /v/, the items are considered as cognates. Table 15 shows that Ila, Lenje, Sala Soli and Tonga are more closely related. Their inter-relatedness may be realized from language contact due to geographical proximity. Mostly the languages under study are languages in the same Bantu Botatwe group and they share some cognates that do not necessarily originate from Proto Bantu.

4.8. Sound Changes

The study has established some phonological changes between the eight languages explored and Proto Bantu. The following table displays the examples of the phonological changes that have taken place between some of the languages under study and Proto Bantu:

Table 15: Examples of Phonological Changes which have taken place between some languages under study and Proto Bantu

Gloss	PB	Lexical item	Phonological Change	Language
Animal	-yàmà	-nyama	[y] → [ɲ] (nazalization)	**Sala**
Bone	-kúpà	-fupa	[k] → [f] (fricativization)	**Nyanja**
Count	-bàd	-bala	[d] → [l] (lateralization)	**Ila**
Male	-dÓmÉ	-lume	[d] → [l] (lateralization)	**Kaonde**
Neck	-kíngÒ	-nsingo	[k] → [ns] (pre-nasalization) (fricativization)	**Tonga**
Name	-ginà	lina	[g] → [l] (lateralization)	**Soli**
Louse	-dá	-nda	[d] → [nd]	**Nyanja**

			(pre-nasalization)	
River	-dÓNgà	-longa	[d] → [l] (lateralization)	**Lenje**
Seed	-bótÒ	mbuto	[b] → [mb] (pre-nasalization)	**Tonga**
Star	-nyÈnyÈdì	nyenyenzi	[d] → [nz] (pre-nasalization)and (fricativization)	**Nyanja**
Swell	-bìmbà	-vimba	[β] → [v] (voicing)	**Kaonde**
Wing	-pàpá	-baba	[p] → [β] (fricativization)	**Sala**

(Source: http://linguistics.berkeley.edu/CBOLD/Docs?Guthrie.html)

The data provided in Table 15 show that there are phonological changes that have taken place within the languages under study. The changes are fricativization, voicing, pre-nasalization and lateralization.

4.9   Summary

Chapter Four dealt with the findings of the study, the criteria used for determining cognates, establishing percentages of inter-relatedness, group average method, cognate retention and shared cognates.

The chapter has shown the degree of inter-relatedness between the languages investigated. It has revealed that Lenje and Tonga have the highest degree of inter-relatedness, which is at 84.5 percent. This finding is a positive one in that it confirms that Lenje learners who take Tonga as a school subject are using a language which is highly related to their own language and with which they are similar. The researcher supports the decision by the

language planners' use of Tonga for initial literacy and as a school subject in Lenje speaking areas due to the high percentage of inter-relatedness between Tonga and Lenje.

This study has also established that Ila and Kaonde, Ila and Nyanja have the lowest degree of inter-relatedness at 30 percent. However, the low percentages of inter-relatedness have no negative impact on these languages because Kaonde children use Kaonde at school and Nyanja is used as a school subject within Lusaka and Eastern provinces. And Ilas use Tonga at school, a language which has been shown to have a high percentage of inter-relatedness with Ila.

Also, the study has revealed that there are cognates that are shared between the languages under study as well as cognates that originate from Proto Bantu. The general observation made was that the Proto Bantu cognates that are found in all the eight languages are numerals, body parts and names of animals. The reasons why numerals, body parts and names of animals are cognates in all the languages investigated and are retained from Proto Bantu forms are not known. Therefore, there is need for further investigations in order to bring out empirical evidence on the matter.

Further, the study has established that the degree of inter-relatedness between Sala and Tonga is at 80 percent. Since Tonga is used for initial literacy in schools within the Sala area, the researcher believes that Sala pupils may not have major problems in learning the Tonga language due to its level of inter-relatedness to Sala.

Further still, this study has established that Soli shares a high percentage of inter-relatedness with Lenje at 72 percent. The high percentage of inter-relatedness between these two languages may be ascribed to the close proximity of the speakers because they

share boundaries between the Great East road and the Great North Road. The investigations at the Curriculum Development Center confirmed that some years ago Lenje was used as a school subject in the Soli speaking areas. Another finding of the study is that the degree of inter-relatedness between Soli and Nyanja is only 45 percent. This finding poses a challenge to language planners in the Soli-speaking Chongwe District because when the percentage of inter-relatedness is low, it is likely that the learners will face challenges in learning a language that is not similar to their own. Interviews conducted by the researcher on 20th February, 2013, in Chongwe with Mr. Jashon Takile, Senior Headman Kakuka, Chief Representative and Chairman for Chakwela Makumbi Ceremony confirmed that the Nyanja taught in schools is different from the Nyanja spoken generally by the people in Chongwe Township. Further, the Chakwela Makumba Ceremony chief representative stated that because of the big difference between Nyanja and Soli, performance in Nyanja is poor for most pupils. Understandably, this is the implication of the low percentage of inter-relatedness between Soli and Nyanja. Therefore, the Ministry of education language planners would do well to review the language policies in Zambia based on the percentage of inter-relatedness between the languages concerned and thre actuallanguage usage.

Finally, this study has shown that there are sound changes that have taken place between the reconstructed ancestor language, Proto Bantu and the eight languages investigated. The following are the phonological changes that have taken place: voicing, fricativization, lateralization, pre-nasalization, vowel raising and vowel fronting. The subsequent chapter will deal with testing for local language performance for Chongwe and Mumbwa district selected schools.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1   Introduction

In this chapter, an overview of the findings is presented in summary form. The chapter includes the findings on the degree of inter-relatedness among the eight languages under study. The languages are: Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga.The chapter also includes the cognate retention of Proto Bantu forms with reference to these languages in summary form. Following the research objectives and questions in this study, conclusions have been reached and recommendations made.

This study involved an investigation into the inter-relatedness of Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga as a contribution to existing linguistic information on Bantu Botatwe languages and other Zambian Languages.

The study tried to establish the degree of inter-relatedness between Sala and other Bantu Botatwe languages namely Ila, Lenje, Soli and Tonga. The major focus of the study was Sala, because it is one of the least documented languages among the Bantu Botatwe Group. The languages that are in contact with Sala, such as Ila, Lenje, and Soli (See Appendix E) have also been included in the study to ascertain how closely related Sala is to the surrounding languages. Kaonde was included in the study due to its relatedness to Kaonde-Ila. Nyanja was also included in the study due to its use in Soli speaking areas. Nyanja is used as the language for initial literacy and as a subject in schools in Chongwe and Kafue Districts where Soli is the indigenous language, especially in the rural areas. Thus, the study tried to determine the degree of inter-relatedness that would help to inform the decision to use Nyanja for initial literacy and as a school subject in Soli speaking areas.

The study also tried to ascertain whether there is any retention of the Proto Bantu vocabulary in the languages under investigation.

The significance of this study was to bridge the gap that exists in linguistic studies especially in determining the degree of inter-relatedness of the languages under review. There was no duplication of the study since none of the previous scholars have ventured into the investigation of the inter-relatedness of the languages concerned (Ila, Kaonde, Lenje, Nyanja, Sala, Soli and Tonga).

The researcher employed both qualitative and quantitative research design methods that required the indigenous speakers of the languages under investigations to give the equivalents of the 200 basic vocabulary items in seven languages. The researcher analyzed the data and presented the results in the form of tables, graphs, and hierarchical tree diagrams. The basic vocabulary included names of animals, body parts, lower numerals, and commonly used verbs and nouns.

Different scholars have used different criteria in determining the cognates among the languages. This study mainly made use of Miti's (1996) criteria which are more inclusive. Other scholars' criteria such as Swadesh (1955), Crowley (1992), Nurse and Philippson (1980) Blount and Curley (1970) and others have been referred to and exemplified in this study. Ethret's (1996) criterion has been utilized as well due to its focus on roots or stems of lexical items to determine cognates. Ethret's (1996) criteria are the best especially in the section that deals with Proto Bantu retained cognates, since the Proto Bantu forms are presented as roots or stems.

Sound correspondences of the lexical items were also taken into consideration, following Nurse and Philippson's (1980), Blount and Curley (1970) criteria, which suggest that lexical items can be identified as cognates, whose difference was only in voice.

5.2   Summary of the Findings

This study has attempted to answer the research questions. As stated earlier,

question (a) was 'what is the degree of inter-relatedness between Sala and other Bantu Botatwe languages namely, Ila, Lenje, Soli and Tonga?' The study has established that the degree of inter-relatedness between Sala and Ila is 70 percent, the degree of inter-relatedness between Sala and Lenje is 81.5 percent while the degree of inter-relatedness between Sala and Soli is 67 percent; Sala and Tonga at 80 percent. This scenario suggests that these languages are mutually intelligible to a large extent. This means that the learners may not have much difficulty in learning in Tonga because the language used at school (Tonga) is closely related to the language they speak as their first language.

Question (b) was 'what is the degree of inter-relatedness between Soli and Nyanja? Nyanja is used for initial literacy and taught as a school subject to Soli children in schools in Soli speaking areas. The study has ascertained that the inter-relatedness between Soli and Nyanja is only 45 percent. This situation may not be conducive especially to the learners of Nyanja who speak Soli as their indigenous language and live far away from the townships where Nyanja is used for communication. The degree of inter-relatedness is low and that may affect their school performance.

The study has established the degree of inter-relatedness between Ila and all the languages under study. The degree of inter-relatedness between Ila and Kaonde is 30 percent while

the degree of inter-relatedness between Ila and Lenje is 66.5 percent. The degree of relatedness between Ila and Nyanja is 30 percent, Ila and Soli is 55 percent while that between Ila and Tonga is 76.5 percent.

The study has revealed that the degree of inter-relatedness between Kaonde and other languages is as follows: Kaonde and Lenje is 41 percent, Kaonde and Nyanja is 32 percent, Kaonde and Soli is 36 percent, while Kaonde and Tonga is at 33 percent (See figure 10 page 93).

It has also been revealed that the degree of inter-relatedness between Lenje and Nyanja is 40 percent, the degree of inter-relatedness between Lenje and Soli is 72 percent, while the degree of inter-relatedness between Lenje and Tonga is 84.5 percent.

The study has also established that the degree of inter-relatedness between Nyanja and Soli is 45 percent while the degree of inter-relatedness between Nyanja and Tonga is 40.5 percent (see Figure 12, page 93). It has been established that the degree of inter-relatedness between Soli and Tonga is 63 percent (see Figure 6, page 85).

Very significantly, the study has revealed that the highest degree of inter-relatedness among the languages under study is between Lenje and Tonga at 84.5 percent (see Figure 6, page 85). The second highest is between Lenje and Sala at 81.5 percent (see Figure 9, page 90) and the third highest is between Tonga and Sala at 80 percent (see Figure 5, page 83), followed by Ila and Tonga at 76.5 percent (see Figure 7, page 87). The lowest percentage is between Ila and Nyanja, both at 30 percent (see Table 4, page 95).

The study has also established the hierarchical group average percentage of inter-relatedness. The group of Tonga and Lenje is at 84.5 percent (see Table 4 page 95), rated

as the highest group followed by Tonga, Lenje and Sala which is at 81 percent (See Table 5, page 97). The fourth group is Tonga, Ila, Lenje and Sala at 70.8 percent (see Table 7, page 98) while Tonga, Ila, Lenje, Sala and Soli group is at 67.3 percent (see Table 8, page 103). The average degree of inter-relatedness of Kaonde, and the rest of the languages apart from Nyanja are at 44.5 (see Table 9, page 104), while Nyanja and the rest of the languages are at 38.1 percent (see Table 10, page 104). The study has established that on the hierarchical group average percentages, Nyanja has the lowest percentage of inter-relatedness to the other seven languages, which is at 38.1 as mentioned above.

Question (c) was 'what is the degree of inter-relatedness between Kaonde and Tonga?' Tonga is used for initial literacy and as a subject in the Kaonde speaking area. The study has established that the inter-relatedness between Kaonde and Tonga is at 33 percent. The low percentage of the inter-relatedness between the two languages may pose a challenge to the Kaonde speaking learners.

Question (d) was 'is there any retention of the Proto-Bantu vocabulary in the languages under investigation? The study of cognate retention from Proto Bantu has shown that out of 200 vocabulary items, 32 items have been found to be cognates in all the eight languages under study, out of which 35 items have been retained from Proto Bantu (PB) in all the languages studied. This signifies the Proto Bantu origins of the stems. In other instances, the cognate's retention involves only two or three languages.

The study has established that the highest percentage of Proto Bantu cognate retention is between Lenje and Tonga with 32 percent, followed by Sala with 30 percent of cognate retention from Proto Bantu stems. Ila has 29 percent of cognate retention from Proto

Bantu stems while Kaonde has 29 percent of cognate retention from Proto Bantu. Nyanja has 28 percent of cognate retention while Soli has the lowest percent of cognate retention from Proto Bantu which stands at 27percent.

While the study has established that there is retention of the Proto Bantu vocabulary in the languages investigated, there are phonological changes that have taken place, for example, fricativization, voicing, pre-nasalization and lateralization. The retention includes numerals such as one, two, three, four, five and ten as shown in Table 10. Some names of animals have also been retained, for example dog and elephant as shown in Table 11. The concept for body parts has also been retained. For example, ear, eye, head, and tooth have been retained from Proto Bantu in all the languages explored apart from 'thigh' which is non-cognate in Kaonde, and Nyanja.

5.3 Conclusion

From the analysis carried out in this study, it is evident that Sala shares between 67-81 percent of inter-relatedness with other Bantu Botatwe languages under study. The highest percentage of inter-relatedness between Sala and the ather three Bantu Botatwe languages is between Sala and Lenje, and is the highest among all the languages under study in relation to Sala.

The study has concluded that Sala children may not have many problems in learning Tonga in schools because of the high degree of relatedness which is at 80 percent.

The study has shown that the percentage of inter-relatedness between Soli and Nyanja is low. The implication is that Nyanja may not be the best regional language for initial literacy and as a school subject, particularly for Soli children.

The group average table has shown that among the seven languages investigated, Tonga and Lenje have the highest degree of inter-relatedness of 84.5 percent and the lowest degree of inter-relatedness is between Nyanja and Ila, and Kaonde and Ila at 30 percent. The implication is that Lenje speakers who use Tonga for literacy and as a school subject may find it easy to learn the language because of the high percentage of relatedness while the Kaonde-Ila speakers may find it difficult to learn Tonga due to the low percentage of inter-relatedness. The information especially on the degree of inter-relatedness would be helpful to language policy makers to reconsider the position of some of the regional languages used in particular areas where the percentage of the degree of inter-relatedness between the languages is low.

The findings of this study have shown that there are relatively few researchers who have carried out studies usinglexicostatistical survey between the seven regional languages and the different indigenous languages used by people in various parts of Zambia. Therefore the placement of the regional languages may cause negative attitude towards these languages by the learners that may result in poor performance.

5.4   Recommendations

From this study, the following recommendations have been drawn:

(a) The researcher recommends that lexicostatistical surveys involving other Zambian languages be carried out, in order to help assess whether the current distribution of the regional languages for initial literacy and as subjects in schools is fairly done in relation to the degree of inter-relatedness between the regional languages and the indigenous languages of people in specific areas.

(b) The researcher recommends that further studies be carried out that will investigate the implications of the low percentage of inter-relatedness between Kaonde and Tonga.

(c) Curriculam Development Center (CDC) needs to develop materials for languages used for initial literacy e.g. Kikaonde in Mumbwa and Soli in Chongwe.

REFERENCES

Administrative Committee ya Lenje Bible Translation Project (2003). *Baibo*. Lusaka: Bible Society of Zambia.

Atkinson, M., Roca, I. M. And Kilby, D. (1982). *Foundations of General Linguistics*. London: Unwin Hymnal Ltd.

Barker, P. (1999). *Mixed Languages*. Paper presented at University of Mancester, USA.

Batibo H. M. (1997). *A Lexicostatistical survey of the Bantu Languages of Botswana*: Unpublished paper presented at University of Botswana, Gaborone,October 1997.

Belmont Report(1979). *National Commission for the Protection of human Subjects of Biomedical and Behavioral Research.*

Blount, B. and Curley, R. (1970). *The Southern Luo Languages: a glottochronological reconstruction,* Journal of African Languages, 9:1-18.

Blust, R. A. (1981). *Variation in the retention rate in Austronesian Languages*. Paper presented at fifth International Conference on Austronesian languages Denpasar, Bali.

Bodmer, F., ed. Lancelot, H. (1972). *The Loom of Language*. New York: W.W. Norton and Co.

Brelsford, W. V. (1956). *Generation of Men: The European Pioneers of Northern Rhodesia*, Steart Manning, Harare (Salisbury).

Bunns, R. B. (2000). *Introduction to Research Methods*. London: SAGE Publications.

Bynon, T. (1977). *Historical Linguistics*. Cambridge: Cambridge University Press.

Campbell, L. (2004). (2nd ed). *Historical Linguistics*: An Introduction. Cambridge: The MIT Press.

Carter, H. (Ed) (2000). *An Outline of Chitonga Grammar*. Lusaka: Printech, Zambia.

Carter, H. (2002). *An Outline of Chitonga Grammar*. Lusaka: Bookworld Publishers.

Carter, H. and D. O'brien (1970). *A Tonga Programmed Course in 70 Lessons.* Unpublished materials produced at the School of Oriental and African Studies, London.

Chanda, V. M. (2008). Study Material, course LAL211 Batch No.1. (unpublished material.)

Chitebeta, A. M. (2007). *The Noun Phrase of Tonga and Lenje: A contrastive study*. Unpublished M. A. Thesis. University of Zambia.

Churchhill Jr. G. A. (1995). *Market Research Methodological Foundations* (6th edition). Forth work: The Dryden Press.

Curriculumn Development Center (2013). National Literacy Framework. Lusaka: Zambia Publishing House.

Cole, D. T. (1969). Bantu Linguistic Studies in South Africa, XXVIII, 3, 1969, pp. 201-210.

Collins, B. (1962). *Tonga Grammar*, London; Longmans, Green & Co.

David, O. And Calvin R. (1977). *Two Studies in Middle American Comparrative Linguistics. Arlington: Summer Institute of Linguistics, University of Texas. Publications in Linguistics, publication number 55.*

Doke, C. M. (1945). *Bantu: Modern Grammatical, Phonetical, and Lexicographical Studies since 1860.* International African Institute. London.

Doke, C.M. (1954). *The Southern Bantu Languages*. London: Oxford University Press.

Dyen, I. (1962)). The *Lexicostatistically Determined Relationship of a Language Group.*International Journal of American Linguistics 28/3.

Dyen, Isidle Kruskal, Joseph and Black, Paul (1992). *An Indo-European Classification, a Lexicostatistical Experiment.* Transactions and American Philosophical Society 82/5.

Embleton, Sheila (1986). *Statistics in Historical Linguistics.*

Ethret, C (1998). *An African Classical Age: Eastern and Southern Africa in World History, 1000 BC to AD 400.* Charlottesville: University of Virginia Press.

Fagan, B. M. (1966). *A Short History of Zambia, from earliest times to until A.D. 1900*, Oxford University Press.

Fortune, G. (1959). *A Preliminary Survey of the Bantu Languages of the Federation.* Lusaka: the Rhodes-Livngstone Institute, University of Zambia.

Fourshey, C. C. (1996). *Southwestern Tanzania: Writing a Precolonial History.* Unpublished Paper, University of Carlifornia at Los Angeles.

Grant, A. P. (2009). Swadesh Centenary Conference, 17-18 January 2009, Max-Plank Institute for Evolutionary Anthropology, Leizing. (Unpublished material)

Greenberg, J.H. (1948). Southwestern Journal of Anthropology.

_____ (2001). *The Methods and Purposes of Linguistic Genetic Classification.*Conference Paper, Center for Advanced Studies in Behavioral Sciences, Stanford.

_____ (1963). *Classification of African Languages*.

Grimm, J. (2010). Deutsche Grammatic , 2nd Ed. John Benjamins Pulishing Company.

Goodman, L. A. (1961). *Snowball Sampling*. Annals of Mathematical Statistics 32:148-70.

Gudschinsky, S. (1956). *The ABC's of Lexicostatistics (Glottochronology)*. In *Word,* 12:175-210.

Guthrie, M. (1948). *The Classification of the African Languages*. London: Oxford University Press.

_____(1970b). Journal of African History, Vol. III, 1962, pp.273-82.

Embleton, S. (1986). *Statistics in Historical Linguistics*. Bochum.

Ethret, (1996). *African Language Association of Southern Africa, Foundation for Education, Science snd Technology*. Indiana: University Press.

Hambaba, J. (2008). *Determination and Modification in English and Tonga.* Unpublished M.A. Thesis. University of Zambia, Lusaka.

Hans, G. M. (1972). *Outline of Lexicostatistical Study of Baque and Mande Languages with a note on Fula (pages 199-209.)*

Heine, H. (1974). *The World Bank Encyclopedia*. Chicago: Field Enterprises 9: 158-59.

Hilary, R. (2009). A *Lexicostatistical Survey of selected North-Eastern and South-Eastern Bantu Languages*. (Unpublished MA Dissertation) University of Botswana.

Hinnebusch, T. (1999). '*Contact and Lexicostatistics in Comparative Bantu Studies*.' In Jean Marie H. and May M. (ed) *Bantu Historical Linguistics*. California: CSLI Publications. 5:-173-203.

Hoijer, H. (1956). *Lexicostatistics: A Critique*. Language 32, 49-60

Hopgood, C.R. (1992). *A Practical Introduction to Chitonga*. Lusaka Zambia Educational Publishing House.

Hymes, Dell (1960). *Lexicostatistics so far*. Current Anthropology, 1/1, 344.

Labroussi, C. (1998). *Le Couloir des Lacs- Contribution Linguistique á l'histoire des populations du sud-ouest de la Tanzanie.* Unpulished PhD dissertation, INALCO, Paris.

Labroussi, C. (1999). *Vowel Systems and Spirintization in Southwest Tanzania,* in Hembert, J-M

and Hyman, L. M. (eds) Bantu Historical Linguistics: Theoratical and Impirical Perspectives. CSLI Publication. Stanford. 335-377.

Lansdown, N. G. (_undated) Indyabuluba Nciitalilili. London: Macmillan & Co. Limited.

Lees, R. B. (1953). *The Basis of Glottochronology. Vol.29*. Linguistic Society of America.

Lehmann, D. (2002). *An Outline of Cinyanja Grammar.*

Luwisha, S. Rev. (1962). *Mukulila Coolwe*. Lusaka: Government Printers.

Madane, A. C. (1908). *Lenje Handbook:* A Short Introduction to the Lenje Dialect Spoken in North-West Rhodesia, London: Oxford University Press.

Mambwe, K. (2008). Some Linguistic variations of Kaonde: A dialectological study. Unpublished M. A. Thesis. University of Zambia.

Manchishi, P. C. and Musona, E. T. (Undated). *The People of Zambia: A Short History of the Soli from 1500 to 1900*. Lusaka: Multimedia Publication.

Masayoshi S. and Thoedora B. Eds. (1996:92). *Approaches to Language Typology*. Martin Haspelmath, Free University of Berlin. Journal of Linguistics, Vol.32, Oxford University Press. (www.oup.com) retrieved on the 10th of January 2012.

Matthews, P. (2005). *The Concise Oxford Dictionary of Linguistics*. New York: Oxford University Press.

Mathias, D. Ed. (2002). *Experimenting in tones: Studies in Science and Language.* Stanford:

Stanford University Press, California.

Meeusen, A. E. (1963). *Morphotonology of Tonga Verb*. Journal of African Languages 11/1:72-92.

Mikaili, M. (1968). *Byaalwile Mushala Mema ke Mali*. Published by the Department of Community Development.

Mildred, S. (2007). *The Methodology of Preliminary outcomes of the manual statistical Analysis Performed on the Indo-European 46 Word List.* (Retrieved and printed on 14 November 2010.)

Ministry of Finance. 2000, *Census Report,* Lusaka: Government Printers.

Ministry of Finance. 2010, *Census Report*, Lusaka: Government Printers.

Miti, L. M. (1979). *Preliminary investigation into relative clause formation in Nsenga*. MA Thesis: University of York. Unpublished material.

_____ (1988). *Tonal Variation in Zambian Chinyanja Varieties: An autosegmental Analysis.* Unpublished PhD Thesis, University of London.

_____ (1996). '*Subgrouping Ngoni varieties within Nguni*: A lexicostatistical approach', South African Journal of African Languages, 16(3).

_____ (2006). *Comparative Bantu Phonology and Morphology*. Cape Town: CASAS.

Mkude, D. J. (1995). *Towards a Semantic Typology of Swahili Language*. Tokyo: FUJIWARA Printing Co. Ltd.

Mugenda, O. M. and Mugenda, A. G (1999). *Research Methods: Quantitative and Qualitative Approaches.* Nairobi: Acts Press.

Mugenda, O. M. and Mugenda, A. G. (2003). *Research Methods: Quantitative and Qualitative Approaches.* Nairobi: Acts Press.

Mutambakyalo, M. (1988). *Kyapusana.* Lusaka: Kenneth Kaunda Doundation.

Mweemba, J.B. (1996). *Mubekwa bekwa.* Lusaka: Zambia Educational Publishing House.

Nchete, D. C. (1964). *Bana Bensiku*. Lusaka: The Publication Bireau, Zambia

Ngalande, S. (2007). *A Logical Analysis of selected texts in Nyanja*. Unpublished M.A Dissertation, Lusaka: University of Zambia.

Nkolola, M. W. (1997). *An analysis of the applied causative and passive verb extensions in Tonga*. Unpublished Thesis (PhD). The University of Zambia.

Nkolola, M. W. (2010). *The Discourse of "Call Boys" and Minibus Conductors in Zambia: a Hybrid Sociolect of Identity*. Journal of Multicultural Discourses 5, 2010 pp. 131-156.

Nurse, D. and Philippson, G. (1980). *The Bantu Languages of East Africa: a lexicostatistical survey.* In Polomé, E.C. and Hill, C. P. (eds.) Language in Tanzania, Oxford: Oxford University Press.

Ohannesian, S. and Kashoki, M. E. (1978). (Eds). *Language in Zambia*. London: International African Institute.

Oliver, R.A. and J. D. Foge (1962). *A Short History of Africa*. Pelican.

Oltrogge, D. And C. R. Rensch (1977). *Lexicostatistical Methods in Comparrative Studies*. Los Angels: University Press.

Paul, V. W. (2005). *Dictionary of Statitics and Methods: A Nontechincal Guide for Social Sciences*. London: Sage.

Phiri, J. E. (1957). Kalenga Ndi Muzace. Cape Town: Oxford University Press.

Postnanski, M. (1966). *Prelude to East African History*. Oxford University Press.

Rensch, H. (1966). *Close Relationship between Subtiaba and Tlappanec of Mexico*. Los Banos: California- Oregon.

Ringe, W. (2005). *Lexical Innovations, Morphological Changes, and Syntactic Changes*. London: Routledge.

Rudestam, K. E. And Newton R. R. (1992). *Sampling in Qualitative Enquiry*. California: New Sage Publication, Inc.

Sibajene, A. (2013). *A Dialectological Study of Tonga*. Unpublished M.A. Dissertation. Lusaka: University of Zambia.

Singleton, R. And B. C. Straits (1999). *Approaches to Social Research*. Oxford: University Press.

Song, J.J. (ed.) (2011). The Oxford Handbook of Linguistic Typology. Oxford: Oxford University Press.

Sudman, S., and G. Kalton (1986). *New Developments in the Sampling of Special Populations.*Annual Review of Sociology 12:401-29.

Swadesh, M. (1950). *Salish International Relationships*. *International Journal of American Linguistics.*16, 157-167.

_____ (1955). *Towards Greater Accuracy in Lexicostatistical Dating. International Journal of American Linguistics.* 21:121-137.

Syamupa (1963). *Bala Umvwe*. Pietermaritsburg: The Nal Witness (PTY) Limited.

Taylor, J. and Bogdan, R. (1984). *Introduction to Qualitative Research Methods*. New York: John Wiley & Sons.

Torrend, J. (1931). *An English Vernacular Dictionary of the Bantu Botatwe Dialects of Northern Rhodesia*. Farnborough: Gragg Press Ltd.

Vanderstoep, S. W. and Johnstone D. D. (2009). *Research Methods for everyday life: blending qualitative and quantitative approaches.* San Francisco: John Wiley and Sons. Inc.

Watch Tower and Tract Society of Pennsylvania, (2011).

Watch Tower and Tract Society of Pennsylvania, (2012).

White, B. (2000). *Dissertation Skills for Business and Management Students*. London and New York: Continuum.

Wright, J. L. (1985). *English-Kaonde Vocabulary. Christian Publishers*: Ndola, Zambia.

Wright, J.L. and Kashoki M. E. (2007). *An Outline of Kikaonde Grammar*. Lusaka: Bookworld Publishing and UNZA Press.

Internet Sources:

htt://linguistics.berkeley.edu/CBOLD/Docs?Guthrie.html, retrieved on 29 February 2012.

http://en.wikipedia.or//wiki/comparative_linguistics#cite_note-1, retrieved 16 June 2011.

http://en.wikitionary.org/wiki:Swadesh_list retrieved on 16 June 2011.

http://en.Wikipedia.org/wiki/Lexicostatics retrieved on 16 June 2011.

http://www.joshuaproject.net/people-profile.php?peo3=12206&rog3=ZA, retrieved on 15 March 2011.

http://www.joshuaproject.net/people-profile.php?peo3=13039&rog3=ZA, retrieved on 15 March 2011.

http://www.joshuaproject.net/people-profile.php?peo3=14977&rog3=ZA, retrieved on 15 March 2011

http://www.panafrifil10n.org/index.php/PanAfrlLoc/ChewaNyanja, retrieved on 16 June 2011

http:/www.people.fas.harvard.edu/~jasnoff/historicallinguistics.html#classical_theory,

retrieved in 21 June 2012.

http://www.wiki.answers.com

Oral Sources:

His Royal Highness Chikumbe Chibamba, Senior Chief Shakumbila. Chisalu Palace: Mumbwa.

Mr. Shachele A. Shibalanga, Chair for the Royal Establishment, Shakumbila, Mumbwa.

His Royal Highness, Chief Mumba, Mumbwa.

Mr. Anderson Mwimbu, Headman, Kambya Mwiimbu Village, Shakumbila, Mumbwa.

Mrs. Rodah Kakoma (senior citizen), Shakumbila, Mumbwa.

Mr. Mundia Mwendende, Principal Curriculum Development Specialist, Languages. Curriculum Development Center, Lusaka.

Mr. J B Mwinga, Principal, Rockview Institute, Mwembeshi, Mumbwa.

Ms. Finess Mbwela, Situmbeko Village, Mumbwa.

Mr. Audry Chaba, Situmbeko Village, Mumbwa.

Mr. Jashon Takile, Senior Headman Kakuka Village, Chief Representative, Chairman for Chakwela Makumbi Ceremony.

Chibwalu Village, Chongwe: Group interviews.

Mrs. R. M. Chulu, Senior Standard Education Officer, Ministry of Education, Livingstone.

APPENDIX A: 200-ITEM WORD LIST TRANSLATED INTO SEVEN LANGUAGES

	ENGLISH	ILA	KIKAONDE	LENJE	NYANJA	SALA	SOLI	CHITONGA
	all	zyo-onse	byo-onse	sho-onse	vo-onse	zho-onse	fyo-onse	zyo-onse
	animal	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama
	ashes	twe	Buto	mulota	phulusa	milota	mulota	Matwe
	back(positio)	kubweela	kunyuma	kusule	mbuyo	kusule	kubweela	Kusule
	bad	cibi	kyatama	cibi	chaipa	cibyabi	chaipa	Cibi
	bark	kukuwa	Kuboza	kukuwa	kukuuwa	kukuwa	kukuwa	Kukuwa
	bathe(v)	ku-samba	ko-vwa	ku-samba	ku-samba	ku-samba	ku-samba	ku-samba
8	belly	da	Jifumo	lifumo	mimba	vumo	libunda	Da
9	Big	cikando	kikatampe	cinene	ci-kulu	ci-nene	chi-nene	ci-nene
10	bird	zuni	Ngonyi	yuni	mbalame	yuni	kakeeni	Yuni
11	bite(v)	ku-luma	ku-sunama	ku-luma	ku-luma	ku-luma	ku-luma	ku-luma

12	black	shiya	Fita	shiya	kuda	shiya	chishipa	Siya
13	blind person	moofu	Mpofu	moofu	khungu	moofu	mpofu	Moofu
14	blood	bulowa	Maashi	bulowa	magazi	bulowa	milopa	Bulowa
15	bone	huwa	Nikupa	fuwa	fupa	cifuwa	chifupa	Fwuwa
16	breast	nkolo	Jibele	lukolo	bere	tukolo	libele	Nkolo
17	breathe(v)	ku-zoya	ku-peema	ku-yoya	ku-puma	ku-yoya	ku-yoya	ku-yoya
18	burn(v)	ku-tenta	ku-sooka	ku-tenta	ku-ocha	ku-tenta	ku-tenta	ku-umpa
19	buy	ku-ula	ku-pota	ku-ula	ku-gula	ku-ula	ku-ula	ku-ula
20	call(v)	koompolola	Kwiita	kwiita	kuitana	kwiita	kukuwa	Kwiita
21	catch(v)	ku-kwata	ku-kwata	ku-cata	ku-gwira	ku-jata	kwi-kata	ku-jata
22	chew(v)	ku-lukuta	ku-nyeuna	ku-tafuna	ku-tafuna	ku-tafuna	ku-tafuna	ku-tafuna
23	child	mwa-nike	mwaana	mwa-nike	mwana	mwana	mwana	Mwana
24	cloud(n)	nkumbi	Jikumbi	nkumbi	mtambo	kunabi	kumbi	Kkumbi

25	cold	tontola	mashika	tontola	zizira	tontola	tontola	Tontola
26	come	hhiza	Iya	kosa	bwera	konza	kwesa	Boola
27	count	ku-bala	ku-belenga	ku-belenga	ku-werenga	ku-bala	ku-belenga	ku-bala
28	cut(v)	ku-kosola	ku-chiba	ku-timbula	ku-dula	ku-zutula	ku-timbula	ku-gonka
29	cry(v)	ku-lila	ku-jila	ku-lila	ku-lira	ku-lila	ku-lila	ku-lila
30	day	buzuba	Juba	bushiku	tsiku	buzuba	busuba	Buzuba
31	die	ku-hwa	ku-taika	ku-fwa	ku-mwalira	ku-fo	ku-fwa	ku-fwa
32	dig	kusya	kupooya	kukaba	kukumba	kusha	kukaba	Kusya
33	dog	mubwa	Kabwa	kabwa	galu	mubwa	kabwa	Mubwa
34	drink(v)	ku-nwa	ku-toma	ku-nwa	ku-mwa	ku-nwa	ku-nwa	ku-nywa
35	dry	Ku-zuma	ku-uma	ku-yuma	ku-uma	ku-yuma	ku-yuma	ci-yumu
36	dumb person	ntaamba	Kibulu	chibuulu	wosatha kulankhula	cibulu	chibulu	Tambi
37	dust(n)	huko	Lukungu	Suko	fumbi	lusuko	lusuko	Suko
38	ear	kutwi	Kutwi	kutwi	khutu	litwi	litwi	Kutwi

39	earth(n)	nshi	Ntanda	chi-shi	dziko lapansi	nyika	chishi	Nyika
40	eat	ku-lya	ku-ja	ku-lya	ku-dya	ku-lya	ku-lya	ku-lya
41	egg	Iyi	Jike	liyi	dzira	liiyi	liyi	Iji
42	eye(n)	liso	Jinso	liinso	diso	linso	liinso	Liso
43	fall(v)	ku-wa	ku-pona	ku-wa	ku-gwa	ku-wa	ku-wa	ku-wa
44	fat(n)	kwiina	Mafuta	Kwiina	mafuta	kwina	kwansuka	Fwuta
45	fear(v)	kuyowa	Kuchina	kuloonda	kuopa	kutinfa	kutina	Kuyoowa
46	fear(n)	bu-kandu	China	boowa	mantha	kutinfa	kutina	Boowa
47	feather	pepe	mayona	lipepe	nthenga	mapepe	mapepe	Mapepe
48	finger	munwe	Munwe	munwe	chala	munwe	ci-mbombo	Munwe
49	fire(n)	mulilo	Mujilo	mulilo	moto	mulilo	mulilo	Mulilo
50	fish(n)	inswi	Jisabi	inswi	nsomba	inswi	inswi	Inswi
51	five	sanwe	Kitanu	sanu	sanu	cisanu	chisanu	Sanwe
52	flower(n)	iluba	Jiluba	liluba	duwa	maluba	liluba	Duba

53	fly(v)	kuluka	kutumbuka	kuluka	kuluka	kuluka	kuuluka	Kuluka
54	foot(n)	humba	Mana	mpata	phazi	cimpata	chimpata	Cituta
55	four	syone	Kiina	shone	nayi	zhone	chinai	Tone
56	frog	bombwe	bombwe	bombwe	chula	bonwe	chula	Cula
57	good	ci-botu	kyawama	Botu	bwino	botu	chaina	Botu
58	give	ku-pa	ku-pana	ku-pa	ku-Patsa	ku-pa	ku-pa	ku-pa
59	go	ku-ya	ku-yanga	ko-ya	ku-pita	ko-ya	ku-ya	ku-unka
60	grass	bwizu	Nsoono	bwiisu	udzu	bwizwu	mwila	Bwizu
61	guts	bohhu	Ngovu	mala	Kulimbika	nguzu	mila	Sicaamba
62	hair	masusu	Nsuki	misusu	tsitsi	misusu	mishishi	Masusu
63	hand(n)	tasyi	Kuboko	lyaansa	dzanja	lyanza	likasa'	Kuboko
64	head(n)	mutwi	Mutwe	mutwi	mutu	mutwi	mutwi	Mutwe
65	hear	teelela	ku-mvwa	nyumfwa	ku-mva	kunva	nyumfwa	Kumvwa
66	heart	monzo	muchima	moyo	mtima	moyo	mooyo	Moyo

67	heavy	ku-lema	ku-nema	ku-lema	ku-lema	ci-lemu	cha-lema	ci-lema
68	here	kono	Pano	Kuno	pano	kuno	pano	Aano
69	horn(n)	lwiya	Isengo	lwiica	nyanga	lwifa	lija	Lwija
70	how?	mbuti?	byepi?	munsilaanshi?	bwanji?	buti?	achoni?	buti?
71	housefly	ihhi	Lonzhi	looshi	nchenche ya myumba	zhimini	shikanshi	Ziniiini
72	kill(v)	ku-yaya	kwi-paya	ku-caya	ku-pha	ku-jaya	ku-shina	ku-jaya
73	know	ku-hhiba	ku-yuka	ku-shiba	ku-dziwa	ku-ono	kwi-nshaba	ku-zyiba
74	laugh(v)	ku-seka	ku-seeka	ku-seka	ku-seka	ku-seka	ku-seka	ku-seka
75	leaf	tuhhu	Jibuula	litewu	tsamba	liteu	litewu	Itu
76	leg	mwendo	Kulu	mwendo	mwendo	mwendo	mwendo	Kulu
77	liver	muni	Jibu	Lini	chiwindi	muni	mpafwa	Muni
78	long(dist ance)	lampa	Kyalepa	lampa	utali	lamfu	kutali	Lamfu
79	louse	jina	Nkuso	njina	inda	njina	njina	njina

80	man/male	mulombwana	wamulume	musankwa	mwamuna	musankwa	mutuloba	mulombwana
81	many	zhinji	Byavula	shinji	mbiri	shinji	bangi	zinji
82	meat	buzani	Nyama	nyama	nyama	nyama	nyama	nyama
83	moon	mwehhi	Nĝondo	mwenshi	mwezi	museezyi	mwenshi	mweezi
84	mountain	lupili	katumba	mulundu	phiri	mulundu	mulondu	cilundu
85	mouth	mulomo	Kanwa	mulomo	kamwa	mulomo	mulomo	mulomo
86	name(n)	ihhina	Jizhina	liina	dzina	zhina	lina	zina
87	neck(n)	ihsyingo	Nshingo	inshingo	khosi	nshingo	inshingo	nsingo
88	night	mashiku	Bufuku	mashiku	usiku	mashiku	mashuku	masiku
89	nose	inango	Moona	nshono	mphuno	mpemo	mupe'mbo'	mpemo
90	now	nihhyo	katataka	ono	tsopano	ono	lino	lino
91	ancient	kalekale	bashakulu	kalekale	kale	kalekale	byakendi	kalekale
92	one	comwi	Kamo	comwi	modzi	comwi	chimo	komwe
93	other	cimwi	kikwabo	cimwi	cina	cimwi	chimbi	zyimwi

94	person	mu-ntu	mu-ntu	mu-ntu	mu-nthu	mu-ntu	mu-ntu	mu-ntu
95	play(v)	ku-sobana	ku-kaya	ku-sekana	ku-sewera	ku-sobana	ku-fukana	ku-sobana
96	pull(v)	ku-kwela	ku-donsa	ku-kwela	ku-koka	ku-kusela	ku-kweela	ku-kwela
97	push(v)	ku-tonka	ku-shinjika	ku-shindika	ku-kankha	ku-shindika	ku-nyaka	ku-tonka
98	rain(n)	ihhula	Mvula	infula	mvula	vula	imfula	mvwula
99	reason(n)	muzeezo	mulangwe	muyeeyo	chifukwa	yeya	kwambeti	muzeezo
100	red	subila	Chila	salala	Fiira	salala	chifubela	salala
101	river	mulonga	Mukola	mulonga	mtsinje	mulonga	mulonga	mulonga
102	road	mukwakwa	mukwakwa	nshila	msewu	zyila	nshila	mugwagwa
103	root(n)	muzauda	muzyazhi	muyanda	muzu	muyanda	muyanda	muyanda
104	rope	lukalo	ntaambo	lutambo	chingwe	lutambo	ntambo	tambo
105	rot(v)	ku-bola	ku-bola	ku-bola	ku-wola	ku-sasa	ku-bola	ku-bola
106	salt(n)	mwino	Mukele	mwiinyo	mchere	mucele	muchele	munyo
107	sand(n)	musenga	musenga	museese	mchenga	musenga	museya	musenga
108	see	ku-bona	ku-mona	ku-bona	ku-ona	ku-langa	ku-bona	ku-bona
109	seed(n)	inseke	Nkunwa	imbuto	mbewu	imbuto	mbewu	mbuto

110	sew(v)	ku-sasa	ku-soona	ku-tunga	ku-soka	ku-tunga	ku-tunga	ku-suma
111	sharp	bosya	Twa	citende	kuthwa	bosya	chalumata	bosya
112	shoulder	ikukko	Nipuzhi	cifushi	phewa	wezo	chifunshi	ikukko
113	short(height)	fwaafwi	Bwipi	fwafwi	fupi	fafi	nipepi	fwaafwi
114	shout(v)	ku-saba	ku-punda	ku-olobesha	ku-fuula	ku-ompolola	ku-shauta	ku-ompolola
115	shut	ku-yala	ku-shinka	ku-cala	ku-tseka	ku-cala	ku-chala	ku-jala
116	sick(be)	kusata	ku-beela	ku-ciswa	ku-dwala	ku-ciswa	ku-kukolwa	ku-ciswa
117	sing	kwimba	Kwimba	kwimba	ku-imba	kwimba	kwi-mbila	kwimba
118	Sit	ku-kkala	Kwikala	kwikala	ku-khala	ku-kala	kwikala	ku-kkala
119	skin(n)	lukanda	Kiseba	cikanda	chikopa	cikanda	chipaya	cikanda
120	sky	izeulu	Jiulu	liculu	mulengalenga	julu	kwilu'	julu
121	sleep(v)	ku-ona	ku-lala	ku-ona	ku-gona	ku-ona	koona	ku-ona
122	small	syoonto	kacheche	kaniini	chepa	shoto	kangana	shonto
123	smell(v)	ku-nunka	ku-nuunka	ku-nunsha	ku-nunkha	ku-nunka	kununka	u-nunka
124	smoke(n)	busyi	Bwishi	bwiishi	utsi	buishi	bwishi	busi
125	snake	inzoka	Muloolo	nsoka	njoka	nzoka	injoka	nzoka

126	speak	kwaamba	Kwamba	kwamba	ku-lankhula	kwamba	kwamba	kwamba
127	spit(v)	ku-hwila	ku-shipa	ku-shipa	ku-Lavula	ku-shimpila	ku-sanka	ku-swida
128	split(v)	kwandanya	kwabanya	ku-pasula	kung'amba	ku-keta	ku-yambanisha	kwandanya
129	soil(n)	bulongo	Mushiji	bulongo	dothi	bulongo	bulongo	bulongo
130	squeeze(v)	ku-syankanya	ku-fiina	ku-tyana	ku-finya	ku-shaukanya	ku-shanta	ku-tyanka
131	stab(v)	ku-yasa	Kwasa	ku-yasa	ku-lansa	ku-yasa	ku-yasa	ku-yasa
132	stand(v)	ku-hhima	kwimana	kwimikana	kwima	kwima	kwiimana	nyamuka
133	star	intongwehhi	kababanga	nyenyeshi	nyenyezi	nyenyeshi	inyenyenshi	nyenyezi
134	stone(n)	ibwe	Jibwe	libwe	mwala	libwe	libwe	Ibbwe
135	steal	kwiba	Iba	kwiipa	Kuba	bba	kwimba	Kubba
136	strength	insana	Ngovu	nkusu	mphamvu	nguzu	ngofu	nguzu
137	suck(v)	ku-nonka	kwamwa	ku-yamwa	ku-yamwa	ku-nyonka	ku-yamwa	ku-nyonka
138	swallow	ku-mina	ku-mima	ku-mina	ku-meza	ku-mina	ku-kafupi	ku-mena
139	sweat(n)	ibee	Jisulwila	nkasaalo	thukuta	nkasalo	inkasalo	nkasalo
140	swell	ku-hhimba	ku-vimba	ku-shimba	ku-tupa	ku-zhimba	ku-shimba	ku-zimba
141	swim(v)	ku-samba	Kowa	ku-samba	ku-sambira	ku-samba	ku-nyanya	ku-yamba

142	tail(n)	mucila	Mukila	mucila	mchira	mucila	muchila	muchila
143	tell(v)	ku-syimwina	Kwamba	ku-lwiita	ku-uza	kwambila	kwambila	kwambila
144	ten	kumi	Jiikumi	Kumi	khumi	kumi	kwakumi	kumi
145	thief	mutewu	kabwalala	mwiipi	mbala	kabwalala	kabwalala	mubbi
146	thigh	ku-belo	ki-joma	ku-belo	nchafu	li-belo	li-belo	ku-belo
147	thin	ku-koka	ku-nyana	ku-koka	ku-wonda	ku-koka	ku-koka	ku-koka
148	think	ku-tweluka	ku-languluka	ku-yeeya	ku-ganiza	ku-yeya	Ku-yeya	ku-yeya
149	thirst(n)	inyota	Nilaka	nyotwa	ludzu	nyotwa	inyotwa	nyota
150	this	cecci	Akye	Ici	ici	ici	ichi	eci
151	thorn	bwiya	Mwiba	muumfwa	munga	muva	munga	mamvwa
152	three	zyotatwe	Tusatu	shotatwe	tatu	tatwe	chitatu	totatwe
153	throw(v)	ku-waala	ku-taya	ku-waala	ku-ponya	ku-sowa	ku-wala	ku-sowa
154	tie(v)	kwanga	ku-kasa	kwanga	ku-manga	ku-anga	ku-sunga	kwanga
155	time	ciindi	Kimye	ciindi	nthawi	ciindi	chindi	chindi
156	take	ku-bweza	ku-seenda	ku-bwesa	ku-tenga	ku-bweza	ku-manta	ku-bweza
157	taste(v)	ku-labila	ku-umvwa	ku-nyumfwa	ku-lawa	ku-solela	ku-laba	ku-labila
158	tear(v)	ku-zapula	ku-mipolo	ku-kwamuna	ku-ng'amba	ku-kwamuna	ku-twamuna	ku-zaula

159	thatch(v)	ku-hhumba	ku-vweta	Ku-fumba	ku-folera	ku-vumba	ku-shimba	ku-vumba
160	today	sunu	Leelo	Sunu	lero	sanu	lelo	sunu
161	tongue	mulakka	Lujimi	mulaka	lilime	mulaka	mulaka	mulaka
162	tooth	lino	Jiino	Liino	dzino	lino	lino	lino
163	touch(v)	ku-kwata	ku-kwata	ku-cata	ku-khuza	ku-jata	ku-kumya	ku-jata
164	tree	ci-samu	Kichi	ci-samu	mtengo	ci-samu	chi-tondo	ci-samu
165	trouble(v)	pehhi	lu-katazho	penshi	vuto	ku-katazhingwa	mapensho	penzi
166	try (v)	ku-sola	kweeseka	kweelesha	ku-Yesa	kwelezya	kwelesha	ku-sola
167	two	syobile	Tubiji	syobilo	ziwiri	zyobilo	tubili	tobile
168	type(n)	mu-syobo	mu-tundu	mu-shobo	Mtundu	mu-shobo	mu-shobo	mu-syobo
169	uproot	ku-hhyula	ku-tupula	ku-shula	ku-zula	ku-zyula	ku-funya	ku-jwa
170	untie	kamantanya	Kasulula	angulula	masula	angulula	sungulula	kwangulula
171	urine	minhyu	masukula	mikonso	mkodzo	minsyu	mitundo	minsyu
172	village	muhhi	Muzhi	mushi	mudzi	muzhi	munshi	munzi
173	voice(n)	izwi	Jiwi	Liswi	liwu	lizwi	liswi	Ijwi
174	vomit(v)	ku-luka	ku-lasa	ku-luka	ku-sanza	ku-luka	ku-luka	ku-luka
175	wake up	buka	Buuka	Buka	uka	buka	punduka	buka

176	walk(v)	kweenda	kweenda	kweenda	ku-yenda	kweenda	kwenda	kweenda
177	warm(be)	ku-kasala	ku-kaba	ku-kasaala	funditsa	kasaba	ku-kasala	ku-kasaala
178	want(v)	ku-zanda	ku-keeba	ku-yanda	ku-funa	ku-yanda	ku-nyanda	ku-yanda
179	water(n)	meehhyi	Meema	maanshi	madzi	mazhi	menshi	meenda
180	wear	sama	Vwala	fwala	vala	sama	kafwala	sama
181	waste(v)	ku-sinya	biswaswa	nyonyoola	ononga	nyonyoka	kutaya	nyonyoona
182	well(get)	kabotu	buloongo	kabotu	bwino	kabotu	chena	kabotu
183	wet(v)	ku-teta	ku-zoba	ku-tontola	ku-nyowa	ku-tontola	ku-totoba	ku-teta
184	what?	ccihhi	kika?	cinshi	chiyani	cizhi	chani?	ninzi?
185	when?	ulili?	Jubaka	ndlili?	liti	lili	lilyoni?	lili?
186	White	tuba	Utooka	Tuba	yera	tuba	chituba	tuba
187	wind(n)	muwo	Mwela	muwo	mphepo	muwo	lukupwe	muwo
188	wing(n)	Ibaba	Jiyona	papaminwa	phiko	baba	papamino	ibaba
189	wipe(v)	pukuta	Wamya	kunkumuna	pukuta	pukuta	kunkumuna	pukuta
190	witch(n)	mulohhi	Mulozhi	muloshi	mfiti	mulozhi	muloshi	mulozi
191	witness(n)	kkamboni	Mbonyi	bumboni	mboni	kamboni	kaamboni	kamboni

192	woman/ female	mukkaintu	wamukazhi	mwanakashi	mkazi	mukaintu	mutukashi	mwanakazi
193	Woods	lukkuni	Nkunyi	nkuni	nkhuni	inkiimi	inkuni	nkuni
194	work(v)	sebenza	Mungilo	sebensa	gwira nchito	sebenza	sebensa	sebenza
195	Yawn	asamuka	Mwau	Aula	yasamula	kwaula	mwawu	kulya mwau
196	Year	mwakka	Mwaka	mwaaka	chaka	mwaka	chaka	mwaka
197	Yesterday	uzona	Kasha	Ciilo	dzulo	Ijilo	lilo	jilo
198	Elephant	muzohhu	Nzovu	nsofu	njovu	muzovu	njofu	muzovwu
199	knee(n)	gondo	Jinungo	linungo	bondo	linungo	linungo	gondo
200	sun(n)	izuba	Juba	lisuba	dzuwa	izuba	lisuba	zuba

APPENDIX B: VOCABULARY CORRESONDENCES

	ENGLISH	ILA	KIKAONDE	LENJE	CHINYANJA	SALA	SOLI	CHITONGA
1	All	+	+	+	+	+	+	+
2	Animal	+	+	+	+	+	+	+
3	Ashes	#	+	+	-	+	+	#
4	Back(position)	#	-	+	-	+	#	+
5	Bad	+	-	+	#	+	#	+
6	Bark	+	-	+	+	+	+	+
7	Bathe(v)	+	-	+	+	+	+	+
8	Belly	#	+	+	-	+	#	#
9	Big	-	-	+	-	+	+	+
10	Bird	+	-	+	-	+	+	+
11	Bite(v)	+	+	+	+	+	+	+
12	Black	+	-	+	-	+	-	+
13	Blind person	+	+	+	-	+	+	+

149

14	Blood	+	-	+	-	+	+	+
15	Bone	+	-	+	+	+	#	+
16	Breast	+	#	+	#	+	#	+
17	Breathe(v)	+	#	+	#	+	+	+
18	Burn(v)	+	-	+	-	+	+	-
19	Buy	+	-	+	+	+	+	+
20	Call(v)	-	+	+	+	+	+	+
21	Catch(v)	+	+	+	-	+	+	+
22	Chew(v)	-	+	+	+	+	+	+
23	Child	#	+	#	+	+	+	+
24	Cloud(n)	+	+	+	-	-	+	+
25	Cold	+	-	-	-	+	+	+
26	Come	+	-	+	-	+	+	-
27	Count	+	+	#	#	+	#	+

28	Cut(v)	-	-	+	+	+	+	-
29	Cry(v)	+	+	+	+	+	+	+
30	Day	+	+	#	#	+	+	+
31	Die	+	-	+	-	+	+	+
32	Dig	+	-	#	-	+	#	+
33	Dog	+	+	+	-	+	+	+
34	Drink(v)	+	+	+	+	+	+	+
35	Dry	+	+	+	+	+	+	+
36	Dumb person	#	+	+	-	+	+	#
37	Dust(n)	+	-	+	-	+	+	+
38	Ear	+	+	+	+	+	+	+
39	Earth(n)	#	-	#	-	+	#	+
40	Eat	+	+	+	+	+	+	+
41	Egg	+	-	+	-	+	+	+

42	Eye(n)	+	+	+	+	+	+	+
43	Fall(v)	+	-	+	+	+	+	+
44	Fat(n)	#	+	#	+	#	-	+
45	Fear(v)	+	-	-	-	#	#	+
46	Fear(n)	-	-	+	-	#	#	+
47	Feather	+	-	+	-	+	+	+
48	Finger	+	-	+	-	+	-	+
49	Fire(n)	+	+	+	-	+	+	+
50	Fish(n)	+	-	+	-	+	+	+
51	Five	+	+	+	+	+	+	+
52	Flower(n)	+	+	+	+	+	+	+
53	Fly(v)	+	+	+	+	+	+	+
54	Foot(n)	-	-	+	-	+	+	+
55	Four	+	+	+	#	+	#	+

56	Frog	+	+	+	#	+	#	#
57	Good	+	-	+	-	+	-	+
58	Give	+	-	+	+	+	+	+
59	Go	+	+	+	+	+	+	-
60	Grass	+	-	+	+	+	-	+
61	Guts	+	+	+	-	+	+	-
62	Hair	+	-	+	+	+	+	+
63	Hand(n)	-	-	+	-	+	-	#
64	Head(n)	+	+	+	+	+	+	+
65	Hear	-	+	+	+	+	+	+
66	Heart	+	#	+	#	+	+	+
67	Heavy	+	+	+	+	+	+	+
68	Here	+	#	+	#	+	#	#
69	Horn(n)	+	-	+	-	+	+	+

70	How?	+	-	-	-	+	-	+
71	Housefly	+	+	+	-	+	+	+
72	Kill(v)	+	+	+	-	+	-	+
73	Know	+	-	+	+	-	+	+
74	Laugh(v)	+	+	+	+	+	+	+
75	Leaf	+	-	+	-	+	+	+
76	Leg	+	#	+	+	+	+	#
77	Liver	+	-	+	-	+	-	+
78	Long(distance)	+	+	+	^	#	^	#
79	Louse	+	-	+	-	+	+	+
80	Man/male	#	-	+	-	+	-	#
81	Many	+	-	+	-	+	+	+
82	Meat	-	+	+	+	+	+	+
83	Moon	+	-	+	+	+	+	+

84	Mountain	#	-	+	#	+	+	+
85	Mouth	+	#	+	#	+	+	+
86	Name(n)	+	+	+	+	+	+	+
87	Neck(n)	+	+	+	-	+	+	+
88	Night	+	+	+	+	+	+	+
89	Nose	-	-	+	+	+	+	+
90	Now	-	-	+	+	+	+	+
91	Ancient	+	-	+	+	+	-	+
92	One	+	#	+	-	+	#	+
93	Other	+	-	+	-	+	-	+
94	Person	+	+	+	+	+	+	+
95	Play(v)	+	-	+	-	+	+	+
96	Pull(v)	+	-	+	-	+	+	+
97	Push(v)	#	+	+	+	+	-	#

98	Rain(n)	+	+	+	+	+	+	+
99	Reason(n)	+	-	+	-	+	+	+
100	Red	#	#	+	#	+	+	+
101	River	+	-	+	-	+	+	+
102	Road	#	-	+	-	+	+	#
103	Root(n)	+	-	+	-	+	+	+
104	Rope	-	+	+	-	-	+	+
105	Rot(v)	+	+	+	+	+	+	+
106	Salt(n)	#	+	#	+	+	+	+
107	Sand(n)	+	+	-	+	+	+	+
108	See	+	+	+	+	+	+	+
109	Seed(n)	-	-	+	#	+	#	+
110	Sew(v)	-	-	+	-	+	+	-
111	Sharp	+	#	-	#	+	-	+
112	Shoulder	+	#	#	-	-	#	+
113	Short(hight)	+	#	+	#	+	#	+
114	Shout(v)	-	-	-	+	+	-	-

115	Shut	+	#	+	-	+	+	+
116	Sick(be)	-	#	+	-	+	-	+
117	Sing	+	+	+	+	+	+	+
118	Sit	+	+	+	+	+	+	+
119	Skin(n)	+	#	+	#	+	-	+
120	Sky	+	+	+	-	+	+	+
121	Sleep(v)	+	#	+	+	+	+	+
122	Small	+	#	-	#	+	-	+
123	Smell(v)	+	+	+	+	+	+	+
124	Smoke(n)	+	+	+	+	+	+	+
125	Snake	+	-	+	+	+	+	+
126	Speak	+	+	+	-	+	+	+
127	Spit(v)	-	+	+	-	+	-	-
128	Split(v)	+	+	-	-	-	-	+
129	Soil(n)	+	-	+	-	+	+	+
130	Squeeze(v)	+	-	#	+	+	-	#
131	Stab(v)	+	+	+	+	+	+	+

132	Stand(v)	+	+	+	+	+	+	+
133	Star	-	-	=	+	+	+	+
134	Stone(n)	+	+	+	-	+	+	+
135	Steal	+	+	+	+	+	+	+
136	Strength	-	#	+	#	+	+	+
137	Suck(v)	+	#	#	#	+	#	+
138	Swallow	+	+	+	+	+	+	+
139	Sweat(n)	-	-	+	-	+	+	+
140	Swell	+	+	+	-	+	+	+
141	Swim(v)	+	-	+	+	+	-	+
142	Tail(n)	+	+	+	+	+	+	+
143	Tell(v)	-	+	-	-	+	+	+
144	Ten	+	+	+	+	+	+	+
145	Thief	-	+	#	-	+	+	#
146	Thigh	+	-	+	-	+	+	+
147	Thin	+	-	+	-	+	+	+
148	Think	-	-	+	-	+	+	+

149	Thirst(n)	+	-	+	-	+	+	+
150	This	+	-	+	+	+	+	+
151	Thorn	-	-	+	#	+	#	+
152	Three	+	+	+	+	+	+	+
153	Throw(v)	+	-	+	-	#	+	#
154	Tie(v)	+	-	+	+	+	+	+
155	Time	+	-	+	-	+	+	+
156	Take	+	-	+	-	+	-	+
157	Taste(v)	+	#	#	+	-	+	+
158	Tear(v)	+	-	+	-	+	+	+
159	Thatch(v)	+	-	+	-	+	+	+
160	Today	+	#	+	#	+	#	+
161	Tongue	+	#	+	#	+	+	+
162	Tooth	+	+	+	+	+	+	+
163	Touch(v)	+	+	+	-	+	-	+
164	Tree	+	#	+	-	+	-	+
165	Trouble(v)	+	-	+	-	-	+	+

166	Try	#	+	+	+	+	+	#
167	Two	+	+	+	+	+	+	+
168	Type(n)	+	#	+	#	+	+	+
169	Uproot	+	-	+	+	+	-	-
170	Untie	-	+	+	+	+	+	+
171	Urine	+	-	#	#	+	-	+
172	Village	+	+	+	+	+	+	+
173	Voice(n)	+	#	+	#	+	+	+
174	Vomit(v)	+	-	+	-	+	+	+
175	Wake up	+	+	+	+	+	+	+
176	Walk(v)	+	+	+	+	+	+	+
177	Warm(be)	+	#	+	-	#	+	+
178	Want(v)	+	-	+	-	+	+	+
179	Water(n)	+	-	+	+	+	+	#
180	Wear	#	+	+	+	#	+	#
181	Waste(v)	-	-	#	-	#	-	+
182	Well(get)	+	-	+	-	+	-	+

183	Wet(v)	#	-	+	-	+	+	#
184	What?	+	-	+	-	+	-	+
185	When?	+	-	+	+	+	-	+
186	White	+	-	+	-	+	-	+
187	Wind(n)	+	-	+	-	+	-	+
188	Wing(n)	+	-	#	-	+	#	+
189	Wipe(v)	+	-	#	+	+	#	+
190	Witch(n)	+	+	+	-	+	+	+
191	Witness(n)	+	+	+	+	+	+	+
192	Woman/ Female	#	+	+	+	#	+	+
193	Woods	+	+	+	+	+	+	+
194	Work(v)	+	-	+	+	+	+	+
195	Yawn	+	+	+	+	+	+	+
196	Year	+	+	+	+	+	+	+
197	Yesterday	-	-	+	+	+	+	+

198	Elephant	+	+	+	+	+	+	+
199	Knee(n)	#	+	+	#	+	+	#
200	Sun(n)	+	+	+	+	+	+	+

APPENDIX C: COGNATES RETAINED FROM PROTO BANTU

PROTO BANTU	ENGLISH	ILA	KIKAONDE	KAONDE ILA	LENJE	CHINYA NJA	SALA	SOLI	CHITONGA
-yonce	All	zyo-onse	byo-onse	byo-onse	sho-onse	vo-onse	zho-onse	fyo-onse	zyo-onse
-yàmà	Animal	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama	mu-nyama
-béép-	Bad	cibi	kyatama	cabipa	cibi	chaipa	cibyabi	chaipa	cibi
-yÓg	Bathe(v)	ku-samba	ko-vwa	koowa	ku-samba	ku-samba	ku-samba	ku-samba	ku-samba
-dóm-	Bite(v)	ku-luma	ku-sunama	ku-suma	ku-luma	ku-luma	ku-luma	ku-luma	ku-luma
-yídò	Black	shiya	fita	fita	shiya	kuda	shiya	chishipa	siya
-pÒkù	Blind person	moofu	mpofu	mpofu	moofu	khungu	moofu	mpofu	moofu
-gàdì	Blood	bulowa	maashi	maashi	bulowa	magazi	lowa	milopa	bulowa
-kúpà	Bone	huwa	nikupa	chikupa	fuwa	fupa	cifuwa	chifupa	kwuwa

-bÉÉdÈ	Breast	nkolo	jibele	jibele	lukolo	bere	tukolo	libele	nkolo
-pÈÈm-	Breathe(v)	ku-zoya	ku-peema	ku-pema	ku-yoya	ku-puma	ku-yoya	ku-yoya	ku-yoya
-gòd	Buy	ku-ula	ku-pota	ku-pota	ku-ula	ku-gula	ku-ula	ku-ula	ku-ula
-takun	Chew(v)	ku-lukuta	ku-nyeuna	ku-nyauna	ku-tafuna	ku-tafuna	ku-tafuna	ku-tafuna	ku-tafuna
-ana	Child	mwa-nike	mwaana	mwaana	mwa-nike	mwana	mwana	mwana	mwana
-dùndÈ	Cloud(n)	nkumbi	jikumbi	kumbi	nkumbi	mtambo	kunabi	kumbi	kkumbi
-píÒ	Cold	tontola	mashika	talaa	tontola	zizira	tontola	tontola	Tontola
-yìj	Come	hhiza	iya	iya	kosa	bwera	konza	kwesa	boola
-bàd	Count	ku-bala	ku-belenga	ku-belenga	ku-belenga	ku-werenga	ku-bala	ku-belenga	ku-bala
-cÉNg or -kèd	Cut(v)	ku-kosola	ku-chiba	ku-chiba	ku-timbula	ku-dula	ku-zutula	ku-timbula	ku-gonka
-did or -dèd	Cry(v)	ku-lila	ku-jila	ku-jila	ku-lila	ku-lira	ku-lila	ku-lila	ku-lila

-kú-	Die	ku-hwa	ku-taika	ku-fwa	ku-fwa	ku-mwalira	ku-fo	ku-fwa	ku-fwa
-tím-	Dig	kusya	kupooya	kupoya	kukaba	kukumba	kusha	kukaba	kusya
-bóà	Dog	mubwa	kabwa	kabwa	kabwa	galu	mubwa	kabwa	mubwa
-nu, nyu	Drink(v)	ku-nwa	ku-toma	ku-tomwa	ku-nwa	ku-mwa	ku-nwa	ku-nwa	ku-nywa
-kám-	Dry	kuzuma	kuma	kuuma	kuyuma	kuuma	kuyuma	kuyuma	ciyumuzyi
-kòNgó	Dust(n)	huko	lukungu	lukungu	suko	fumbi	lusuko	lusuko	suko
-tó	Ear	kutwi	kutwi	kutwi	kutwi	khutu	litwi	litwi	kutwi
-dí -dé-	Eat	ku-lya	ku-ja	ku-ja	ku-lya	ku-dya	ku-lya	ku-lya	ku-lya
-jogu	Elephant	muzohhu	nzovu	nzovu	nsofu	njovu	muzovu	njofu	muzovwu
-gé	Egg	Iyi	Jike	jike	liyi	dzira	liiyi	liyi	iji
-yico	Eye(n)	liso	jinso	jinso	liinso	diso	linso	liinso	liso
-nòÈ	Finger	munwe	munwe	munwe	munwe	chala	munwe	ci-mbombo	munwe
-péà	Fire(n)	mulilo	mujilo	mujilo	mulilo	moto	mulilo	mulilo	mulilo

-cú	Fish(n)	inswi	Jisabi	jisabi	inswi	nsomba	inswi	inswi	inswi
-caanu, -taanu	Five	sanwe	Kitanu	chitanu	Sanu	Sanu	cisanu	chisanu	sanwe
-gì	Fly(v)	kuluka	kutumbuka	kuluka	kuluka	kuluka	kuluka	kuuluka	kuluka
-na, -ne	Four	syone	Kiina	tuna	shone	nayi	zhone	chinai	tone
-yòdá	Frog	bombwe	bombwe	bombwe	bombwe	chula	bonwe	chula	cula
-Pá-	Give	ku-pa	ku-pana	ku-pa	ku-pa	ku-patsa	ku-pa	ku-pa	ku-pa
-gè-	Go	ku-ya	ku-yanga	ku-ya	ko-ya	ku-pita	ko-ya	ku-ya	ku-unka
-bÒkÓ	Hand(n)	tasyi	kuboko	kuboko	lyaansa	dzanja	lyanza	likasa'	kuboko
-tó	Head(n)	mutwi	mutwe	mutwe	mutwi	mutu	mutwi	mutwi	mutwe
-yígù	Hear	teelela	ku-mvwa	ku-mvwa	nyumfwa	ku-mva	kunva	nyumfwa	kumvwa
-témà	Heart	monzo	muchima	muchima	moyo	mtima	moyo	mooyo	moyo
-yìpóg-	Kill(v)	ku-yaya	kwi-paya	kwi-paya	ku-caya	ku-pha	ku-jaya	ku-shina	ku-jaya
-cek	Laugh(v)	ku-seka	ku-seeka	ku-seka	ku-seka	ku-seka	ku-seka	ku-seka	ku-seka
-yánì	Leaf	tuhhu	jibuula	jibuula	litewu	tsamba	liteu	litewu	itu

-gòdò	Leg	mwendo	kulu	kuulu	mwendo	mwendo	mwendo	mwendo	kulu
-dà	Long(distance)	lampa	kyalepa	chalepa	lampa	utali	lamfu	kutali	lamfu
-dá	Louse	jina	nkuso	nkuso	njina	Inda	njina	njina	njina
-dómÈ	Man/male	mulombwana	wamulume	mwanamulume	musankwa	mwamuna	musankwa	mutuloba	mulombwana
-yìNgé	Many	zhinji	byavula	byavula	shinji	mbiri	shinji	bangi	zinji
-yàmà	Meat	buzani	nyama	nyama	nyama	nyama	nyama	nyama	nyama
-gìnà	Name(n)	ihhina	jizhina	jizhina	liina	dzina	zhina	lina	zina
-kíNgÒ	Neck(n)	ihsyingo	Nshingo	nshingo	inshingo	khosi	nshingo	inshingo	nsingo
-tíkò	Night	mashiku	bufuku	bufuku	mashiku	usiku	mashiku	mashuku	masiku
-jódò	Nose	inango	moona	mona	nshono	mphuno	mpemo	mupe'mbo'	mpemo
-mó	One	comwi	kamo	kamo	comwi	modzi	comwi	chimo	komwe
-ntu	Person	mu-ntu	mu-ntu	mu-ntu	mu-ntu	mu-nthu	mu-ntu	mu-ntu	mu-ntu

167

Dìòd	Pull(v)	ku-kwela	ku-donsa	ku-doonsa	ku-kwela	ku-koka	ku-kusela	ku-kweela	ku-kwela
-tínd	Push(v)	ku-tonka	ku-shinjika	ku-shindika	ku-shindika	ku-kankha	ku-shindika	ku-nyaka	ku-tonka
-buda	Rain(n)	ihhula	mvula	mvula	infula	mvula	vula	imfula	mvwula
-dÒNgà	River	mulonga	mukola	mukola	mulonga	mtsinje	mulonga	mulonga	mulonga
-jida	Road/path	mukwakwa	mukwakwa	jishinda	nshila	msewu	zhila	nshila	mugwagwa/ nzila
-dì	Root(n)	muzauda	muzyazhi	muzyazhi	muyanda	muzu	muyanda	muyanda	muyanda
-bÒd-	Rot(v)	ku-bola	ku-bola	ku-bola	ku-bola	ku-wla	ku-sasa	ku-bola	ku-bola
-cÉkÉ	Sand(n)	musenga	Musenga	santi	museese	mchenga	musenga	museya	musenga
-bon, -mon	See	ku-bona	ku-mona	ku-mona	ku-bona	ku-ona	ku-langa	ku-bona	ku-bona
-bòtÒ	Seed(n)	inseke	nkunwa	mbuto	imbuto	mbewu	imbuto	mbewu	mbuto
-túm	Sew(v)	ku-sasa	ku-soona	ku-tunga	ku-tunga	ku-soka	ku-tunga	ku-tunga	ku-suma
-túùdí	Shoulder	ikukko	nipuzhi	Chipeya	cifushi	phewa	wezo	chifunshi	ikukko

-kúpé	Short(height)	fwaafwi	bwipi	kuipipa	fwafwi	fupi	fafi	nipepi	fwaafwi
-pám-	Shout(v)	ku-saba	ku-punda	ku-fooloma	ku-olobesha	ku-fuula	ku-ompolola	ku-shauta	ku-ompolola
-dìb-	Shut	ku-yala	ku-shinka	ku-shinka	ku-cala	ku-tseka	ku-cala	ku-chala	ku-jala
-kad	Sit	ku-kkala	kwikala	kwikala	kwikala	ku-khala	ku-kala	kwikala	ku-kkala
-còb-	Skin(n)	lukanda	kiseba	chikoba	cikanda	chikopa	cikanda	chipaya	cikanda
-gòdò	Sky	izeulu	jiulu	jiulu	liculu	mulengalenga	Julu	kwilu'	julu
Dáád-	Sleep(v)	ku-ona	ku-lala	ku-lala	ku-ona	ku-gona	ku-ona	koona	ku-ona
-nîìní	Small	syoonto	Kacheche	kacheche	kaniini	chepa	shoto	kangana	shonto
-NùNK-	Smell(v)	ku-nunka	ku-nuunka	ku-nunka	ku-nunsha	ku-nunkha	ku-nunka	kununka	u-nunka
-yoki	Smoke(n)	busyi	bwishi	bwishi	bwiishi	utsi	buishi	bwishi	busi
-joka	Snake	inzoka	muloolo	mulolo	nsoka	njoka	nzoka	injoka	nzoka

-gàmb-	Speak	kwaamba	kwamba	kwamba	kwamba	ku-lankhula	kwamba	kwamba	kwamba
-tú-	Spit(v)	ku-hwila	ku-shipa	ku-shipa	ku-shipa	ku-Lavula	ku-shimpila	ku-sanka	ku-swida
-bú	Soil(n)	bulongo	mushiji	maloba	bulongo	dothi	bulongo	bulongo	bulongo
-pìn-	Squeeze(v)	ku-syankanya	ku-fiina	ku-shishinkana	ku-tyana	ku-finya	ku-shaukanya	ku-shanta	ku-tyanka
-nyÉnyÈdí	Star	intongwehhi	kababanga	kabangabanga	nyenyeshi	nyenyezi	nyenyeshi	inyenyenshi	nyenyezi
-b ÒÈ	Stone(n)	ibwe	jibwe	jibwe	libwe	mwala	libwe	libwe	ibbwe
-yíb	Steal	kwiba	iba	kwiba	kwiipa	kuba	bba	kwimba	kubba
-med -mid	Swallow	ku-mina	ku-mima	ku-mina	ku-mina	ku-meza	ku-mina	ku-kafupi	ku-mena
Bìmbà	Swell	ku-hhimba	ku-vimba	ku-vimba	ku-shimba	ku-tupa	ku-zhimba	ku-shimba	ku-zimba
-kèdà	Tail(n)	mucila	mukila	mucila	mucila	mchira	mucila	muchila	muchila

-kume -kumi	Ten	kumi	jii-kumi	ji-kumi	kumi	khumi	kumi	kwa-kumi	kumi
-yíbì	Thief	mutewu	Kabwalala	kabwalala	mwiipi	mbala	kabwalala	kabwalala	mubbi
-bÈdÒ	Thigh	ku-belo	ki-joma	chi-jooma	ku-belo	nchafu	li-belo	li-belo	ku-belo
-gàn-	Think	ku-tweluka	ku-languluka	ku-languluka	ku-yeeya	ku-ganiza	ku-yeya	ku-yeya	ku-yeya
-yÓtà	Thirst(n)	inyota	nilaka	chilaka	nyotwa	ludzu	nyotwa	inyotwa	nyota
-yíbà	Thorn	bwiya	mwiba	miba	muumfwa	munga	muva	munga	mamvwa
-catu, -tatu	Three	zyo-tatwe	tu-satu	bi-satu	sho-tatwe	tatu	tatwe	chi-tatu	to-tatwe
-tág-	Throw(v)	ku-waala	ku-taya	ku-taya	ku-waala	ku-ponya	ku-sowa	ku-wala	ku-sowa
-bándod-	Tear(v)	ku-zapula	ku-mipolo	ku-tabula	ku-kwamuna	ku-ng'amba	ku-kwamuna	ku-twamuna	ku-zaula
-dÈÈdÓ-	Today	sunu	leelo	lelo	Sunu	lero	sanu	lelo	sunu

-yínÒ	Tooth	lino	jiino	jino	Liino	dzino	lino	lino	lino
-bidi	Two	syobile	jubiji	tubiji	syobilo	ziwiri	zhobilo	tubili	tobile
-bunga, -ji,-gi	Village	muhhi	muzhi	muzhi	mushi	mudzi	muzhi	munshi	munzi
-jóì	Voice(n)	izwi	jiwi	jiwi	Liswi	liwu	lizwi	liswi	ijwi
-dók-	Vomit(v)	ku-luka	ku-lasa	ku-lasa	ku-luka	ku-sanza	ku-luka	ku-luka	ku-luka
-gÈnd-	Walk(v)	kweenda	kweenda	kweenda	kweenda	ku-yenda	kweenda	kwenda	kweenda
-dúád-	Wear	sama	vwala	vwala	fwala	vala	sama	kafwala	sama
-yÉdó	White	tuba	utooka	chitoka	Tuba	year	tuba	chituba	tuba
-pÉpÒ	Wind(n)	muwo	mwela	luvula	muwo	mphepo	muwo	lukupwe	muwo
-pàpá	Wing(n)	ibaba	jiyona	kapapatiko	papaminwa	Phiko	baba	papamino	ibaba
-pé	Witch(n)	mulohhi	mulozhi	mulozhi	muloshi	mfiti	mulozhi	muloshi	bulozi
-kádì	Woman/ Female	mukkaintu	wamukazhi	mwanamukazhi	mwanakashi	mkazi	mukaintu	mutukashi	mwanakazi
-yáyò	Yawn	asamuka	mwau	mwau	Aula	yasamula	kwaula	mwawu	kulya

									myau
-yákà	Year	mwakka	mwaka	mwaka	mwaaka	chaka	mwaka	chaka	mwaka
-dù	Knee(n)	gondo	inungo	jinungo	linungo	bondo	linungo	linungo	gondo
-jóbà	Sun(n)	izuba	uba	juba	lisuba	dzuwa	izuba	lisuba	zuba

Source: htt://linguistics.berkeley.edu/CBOLD/Docs?Guthrie.html.

Retrieved: 29 February 2012

Appendix D: Map showing Sala Settlement

Scale 1—1500000

10   10   20   20   40 KM

Appendix E: Zambian Map Showing Languages under Study
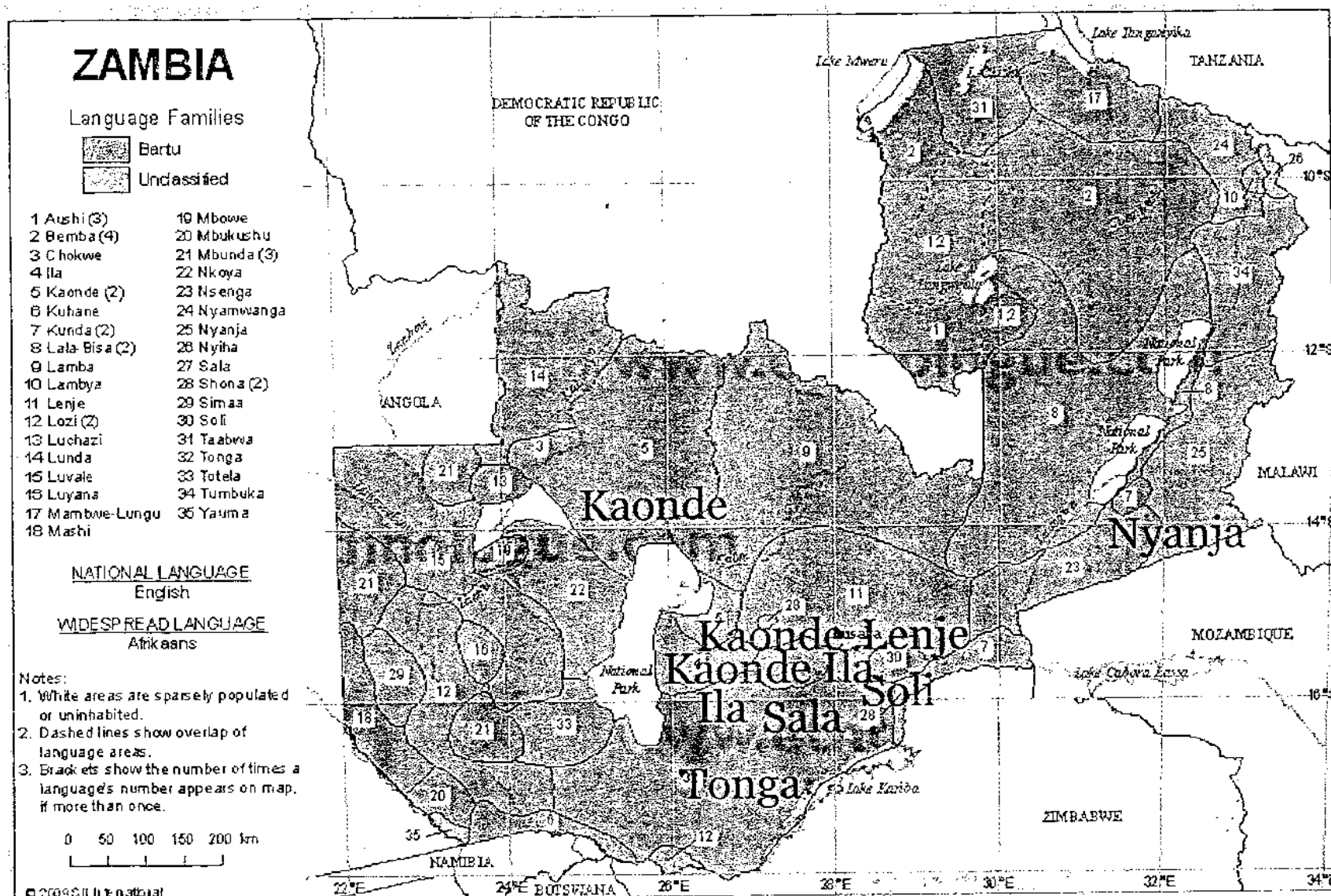
Appendix F: Widely used languages of Communication by Province, Zambia 2010

Language of Communication	Total	central	Copper-belt	Eastern	Luapula	Lusaka	muchinga	Northern	North Western	Southern	Western
Bemba	33.5	31.8	83.9	0.6	71.3	17.6	46.9	69.2	4.9	2.8	0.5
Lala	1.8	17.3	0.3	0.1	0.1	0.1	0.4	0.0	0.0	0.0	0.0
Bisa	1.0	0.0	0.0	1.0	0.1	0.0	6.4	6.2	0.0	0.0	0.0
Ushi	0.9	0.0	0.1	0.0	11.8	0.0	0.0	0.0	0.0	0.0	0.0
Lamba	1.8	2.3	9.2	0.0	0.0	0.1	0.0	0.0	1.7	0.0	0.0
Tonga	11.4	15.5	0.8	0.1	0.1	4.3	0.1	0.0	0.3	74.7	0.2
Lenje	1.2	10.4	0.1	0.0	0.0	0.6	0.0	0.0	0.0	0.1	0.0
Ila	0.7	2.7	0.0	0.0	0.0	0.1	0.0	0.0	0.0	3.7	0.0
Toka Leya	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0
Luvale	1.5	0.2	0.3	0.0	0.0	0.2	0.0	0.0	19.5	0.4	5.1
Lunda(North Western)	1.9	0.1	0.3	0.0	0.0	0.1	0.0	0.0	33.8	0.1	0.3
Mbunda	0.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.2	10.1
Kaonde	1.9	0.9	0.7	0.0	0.0	0.2	0.0	0.0	29.6	0.1	0.3
Lozi	5.5	1.0	0.3	0.0	0.0	1.3	0.0	0.0	0.7	4.0	69.6
Chewa	4.5	0.4	0.1	34.6	0.0	1.2	0.0	0.0	0.0	0.2	0.1
Nsenga	3.0	0.4	0.1	21.4	0.0	1.6	0.0	0.0	0.0	0.2	0.0
Ngoni	0.7	0.3	0.1	4.6	0.0	0.4	0.0	0.0	0.0	0.2	0.0
Nyanja	14.8	8.9	0.7	17.4	0.1	61.9	0.3	0.1	0.6	7.0	0.5
Tumbuka	2.6	0.2	0.2	16.5	0.0	0.4	8.2	0.0	0.0	0.1	0.0
Senga	0.7	0.1	0.0	0.3	0.0	0.1	12.4	0.0	0.0	0.0	0.0
Lungu	0.6	0.0	0.0	0.0	0.0	0.0	0.0	6.9	0.0	0.0	0.0
Mambwe	1.3	0.2	0.1	0.0	0.0	0.3	0.5	14.0	0.0	0.1	0.0
Namwanga	1.2	0.2	0.2	0.0	0.0	0.2	20.7	0.3	0.0	0.0	0.0
English	1.7	0.7	2.1	0.2	0.1	6.2	0.1	0.1	0.6	1.0	0.1
Other Language	4.7	6.1	0.5	3.3	16.3	3.1	3.8	3.0	7.4	1.2	13.2
Percent Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Population	11,126,922	1,098,142	1,741,192	1,349,207	827,639	1,926,022	593,539	918,385	596,860	1,338,649	737,287

Source: 2010 Census of Population and Housing

[i] Cognates are forms that are descendants of the same source. These may be sounds, morphemes or words. For example, the words indicated below from various Bantu languages were derived from the same proto-word. Hence their meaning is the same, i.e. 'person':

(a) isiZulu   umuntu
(b) siSwati   muntfu
(c) ciCewa   munthu
(d) tshiVenda   muthu
(e) sePedi   motho
(f) kiSwahili   mtu
(g) xiTsonga   munhu