Chapter 3

Aim for success, not perfection. Never give up your right to be wrong, because then you will lose the ability to learn new things and move forward with your life. Remember that fear always lurks behind perfectionism. Confronting your fears and allowing yourself the right to be human can, paradoxically, make yourself a happier and more productive person.

Dr. David M. Burns

3 Zulu: a linguistic overview

3.1 Introduction

The topic concerning South African languages can be viewed from different angles. An attempt will be made to establish how the position and use of South African languages, especially Zulu, form part of a broader discourse regarding language, culture and power. However, the primary purpose of this chapter is to create an understanding of the rules of the Zulu language; to analyse the problems and also propose solutions to certain errors experienced during the empirical study described in Chapter 4. The main focus of this chapter is on term creation in Zulu, by:

- means of semantic shifts,
- compounding different words,
- using deideophones or adoptives,
- deverbalisation, and
- loan translation.

By identifying orthographical rules in the Zulu texts, an accurate analysis of errors (in Chapter 5) can be produced.

3.1.1 South African languages after 1994

It is evident that the South African language dispensation has changed quite radically with democratisation in 1994. Before democratisation, languages other than Afrikaans and English did not have the same opportunities for growth and development. This in spite of the fact that approximately 70% of all South Africans have an indigenous language as their mother tongue, whereas 25% have English or Afrikaans as their mother tongue (see Figure 3.1) (Kaschula and De Vries, 2000; Wasserman, 2000).
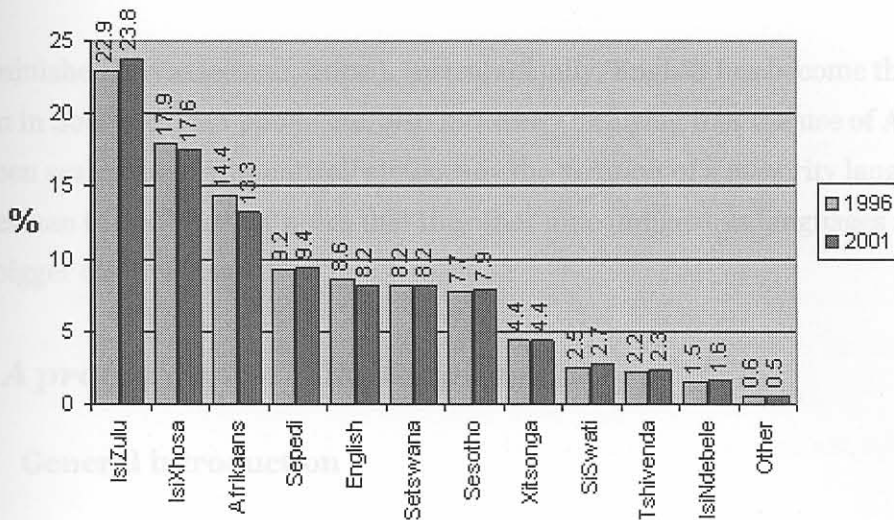
Figure 3.1 *Mothertongue speakers in South Africa* (Statistics South Africa, 2003)

The following table confirms the numbers quoted in Figure 3.1, by stating the figures for each of the language groups since 1990 (while also showing the growth of each language).

Table 3.1 *Number of speakers per language in South Africa (1980, 1991, 1996, 1998 and 2001)*, (Statistics South Africa, 2003).

LANGUAGE	1980	1991 (a)	1996	1998 (b)	2001 (c)
isiZulu	6 064 480	8 343 587	9 200 144	10 194 787	10 677 305
isiXhosa	2 879 360	6 729 281	7 196 118	7 610 435	7 907 153
Afrikaans	4 925 760	5 685 403	5 811 547	5 945 805	5 983 426
Sepedi	2 431 760	n/a	3 695 846	3 832 645	4 208 980
English	2 815 640	3 422 503	3 457 467	3 692 157	3 673 203
Setswana	1 444 908	3 368 544	3 301 774	3 613 925	3 677 016
Sesotho	1 877 840	n/a	3 104 197	3 539 261	3 555 186
Xitsonga	888 140	1 439 809	1 756 105	1 776 505	1 992 207
siSwati	650 600	952 478	1 013 193	1 068 733	1 194 430
Tshivenda	169 740	673 538	876 409	1 227 824	1 021 757
isiNdebele	459 880	n/a	586 961	654 304	711 821
Other	292 360	640 277	228 275	157 767	217 293
Unspecified	n/a	n/a	355 538	10 868	n/a
TOTAL:	26 271 060	31 255 420	40 583 574	43 325 017	44 819 778

The new constitution awarded equal official status (on a national level) to all eleven languages in South Africa. They are: IsiZulu, IsiXhosa, IsiNdebele, Siswati, Sesotho, Setswana, Tshivenda, Afrikaans, English, Xitsonga and Northern Sotho (Sepedi). While this step indicated that indigenous languages are to be developed and their limited rights to be extended, it also stipulated that "existing language rights may not

be diminished" (Wasserman, 2000). Instrumentally, English has become the *lingua franca* in South African public life, also indirectly implying that the use of Afrikaans has been scaled down dramatically to occupy the position of a minority language. Wasserman (2000) further notes that the other nine indigenous languages are at an even bigger disadvantage.

3.2 A profile on the Zulu language

3.2.1 General introduction

More than 10 million people in South Africa has Zulu as their home language (UCLA Language Materials Project, 2001; South African Languages–isiZulu, 2001), with the main concentration of people in Kwazulu-Natal. It is also a dominant language in at least a dozen districts in Gauteng and the Free State. Of all the languages spoken in South Africa, including Afrikaans and English, it has the largest number of speakers– comprising of more than 23% of the total population (ETHNOLOGUE report for language code: zuu, 2001).

3.2.2 Linguistic affiliation

Zulu is a Nguni language that includes Xhosa, Swati and Ndebele. These languages are closely related and mutually intelligible. Nonetheless, they are easily considered dialects of the same language for cultural, historical and political reasons. For instance, Zulu and Xhosa have their own identities according to the individual speakers of the respective languages. The Nguni languages are part of a much larger related group of southern Bantu languages, as reflected in Figure 3.2:
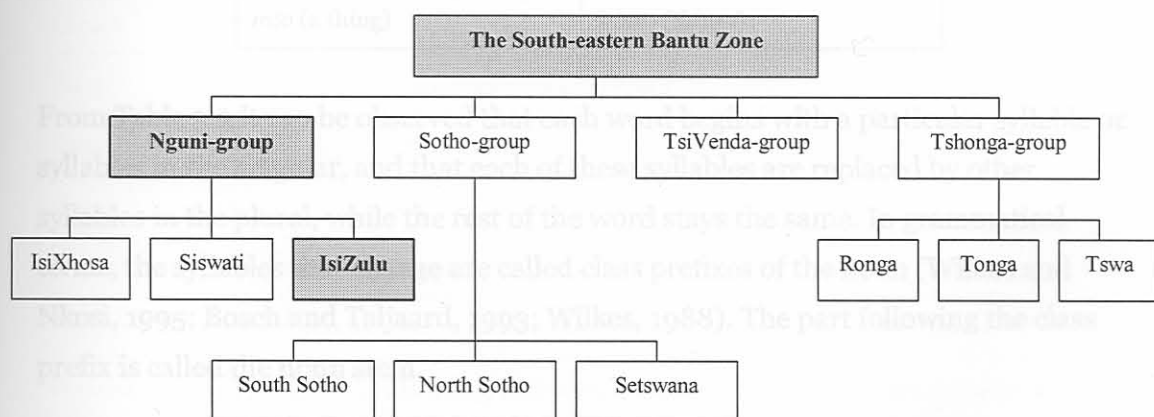
Figure 3.2 The language family tree for Bantu languages

3.3 *The linguistic structure of Zulu*

As Zulu is classified as an agglutinative language, prefixes and suffixes are attached to roots and stems of words, thus conveying the grammatical information. Zulu nouns are divided into sets of classes, called grammatical genders. Each gender has two distinct prefixes, one creating singular nouns and the other plural nouns. The numerous classes far exceed the familiar European classifications (of masculine, feminine and neutral), and are roughly associated with certain semantic characteristics relating to human beings, kinship terms, animals, plants, artefacts, abstract concepts and so on (Cosijn et al., 2002c).

The structure of Zulu is based on two principles (Bosch and Taljaard, 1993; Doke, 1968; Wilkes and Nkosi, 1995): *viz* (a) the system of noun classes, and (b) the system of concords.

3.3.1 The system of noun classes

In Zulu each person or thing, concrete or abstract, is placed in a specific category or group. In grammatical terms, we speak of nouns placed in classes.
Table 3.2 for instance, indicates how nouns are classified into singular and plural forms.

Table 3.2 *Classifying nouns into singular and plural*

Singular	Plural
umuntu (a person)	*abantu* (people)
into (a thing)	*izinto* (things)

From Table 3.2 it can be observed that each word begins with a particular syllable or syllables in the singular, and that each of these syllables are replaced by other syllables in the plural, while the rest of the word stays the same. In grammatical terms, the syllables that change are called class prefixes of the noun (Wilkes and Nkosi, 1995; Bosch and Taljaard, 1993; Wilkes, 1988). The part following the class prefix is called the noun stem.

The stem remains constant, while the prefix may change, for example with the stem -*khulu* the following nouns may be derived:

Table 3.3 Different nouns derived from a stem

ikhulu (a hundred)	*amakhulu* (hundreds)
isikhulu (a person of note)	*izikhulu* (persons of note)
ubukhulu (size)	

From the above it is clear that, although the stem remains the same, the different class prefixes give different semantic contexts to each noun. Furthermore, as Bosch and Taljaard (1993, 1988) notes, these different class prefixes place each one of these nouns into a different noun class, even if the stem remains the same. Today, it is almost impossible for a non-mother tongue speaker to know to which class a certain noun belongs, because the majority of classes contain such a diversity of different nouns. The different noun classes are indicated in Table 3.4.

Table 3.4 The different noun classes

Class	Prefix	Meaning	Example
Class 1	*umu-*	Persons	*umuntu* (person)
Class 1a	*u-*	Terms of relationship, proper names	*ubaba* (my father)
Class 2	*aba-*	Plural of class 1	*abantu* (people)
Class 2a	*o-*	Plural of class 1a	*obaba* (fathers)
Class 3	*umu-*	Natural phenomenon	*umuthi* (tree)
Class 4	*imi-*	Plural of class 3	*imithi* (trees)
Class 5	*ili-*	Miscellaneous	*ilitshe* (a stone)
Class 6	*ama-*	Plural of class 5, collectives	*amatshe* (stones)
Class 7	*isi-*	Implements, miscellaneous	*isitsha* (a plate)
Class 8	*izi-*	Plural class 7	*izitsha* (plates)
Class 9	*in-*	Animals, abstracts, miscellaneous	*inja* (dog)
Class 10	*izin-*	Plural class 9	*izinja* (dogs)
Class 11	*ulu-* or *u-*	Long objects, miscellaneous	*uthi* (stick)
Class 14	*ubu-*	Abstracts, collectives, NO plural	*ubukhulu* (largeness)
Class 15	*uku-*	Infinitives from verb stems	*ukubana* (to see)
Class 16	*pha-*	Used to form locatives (no longer active noun class)	*phandle* (outside)

3.3.2 The system of concords

The class prefix is not only important because it indicates the classes to which the different nouns belong, but also because it links the noun to other words in a sentence (Bosch and Taljaard, 1988, 1993). This is done by means of a concord derived from the class prefix of the noun. This concord is then prefixed to the verb in the sentence.

The system of concordial agreement is important, because it forms the basis of the whole sentence structure of Zulu. For example in the sentence: "The woman washes the plate", the word for "the woman" in Zulu is *inkosikazi*, for 'wash' it is *–geza*, and for 'plate' it is *isitsha*. The subject noun *inkosikazi* must now be brought into concordial agreement with the verb root *–geza* by means of a subject concord *i–* (derived from class prefix *in–*): *Inkosikazi igeza isitsha*.

3.3.3 Other grammatical structures

Not only are concords derived from class prefixes of nouns, but **pronouns** and nominal stems as well. For example:

<center>

Isitsha sona lesi umbala waso muhle.

The colour of this plate is beautiful.

</center>

In the above example, *lesi* is a demonstrative that are found in Class 7 (the isi- class), whereas the pronouns *sona* and *waso*, are formed by means of the following, as indicated in Table 3.5:

Table 3.5 Formation of pronouns

Class	Word (s) referred to	Pronoun stem	Additional prefix/suffix	Pronoun formed
Class 7 (isi-)	isitsha	so-	absolute pronoun (-na)	sona
Class 1 (umu-)	umbala / muhle	-so	possessive concord (wa-)	waso

The **verb** in Zulu has one important characteristic, *viz* that it consists of a root or a radical that carries its basic meaning (Bosch and Taljaard, 1993). However, the root by itself cannot be used, so one needs to add certain prefixes and suffixes for the root to have a functional value.

The different suffixes (also called verbal extensions) are added to the root in order for the verb to be applied in different contexts. When a suffix is added to the root, it becomes a verb stem. For example, the root *FUND* has the basic meaning 'learn', and by suffixing different extensions, it acquires different meanings:

Table 3.6 *Adding verbal extensions (suffixes) to a root give the verb different meanings* (Bosch and Taljaard, 1988)

Verb stem and suffix (Zulu)	Meaning (English)
-*FUNDa*	learn or read
-*FUNDela*	learn for
-*FUNDile*	have learned
-*FUNDisa*	cause to learn/teach/instruct
-*FUNDana*	learn together
-*FUNDisana*	teach one another

As indicated in the above example, the verb is conjugated when represented by the different moods (indicative, imperative and subjunctive), tenses (present, future and past) and other verbal forms in which the verb stem may be used to appear as a complete word.

3.3.4 Term creation in Zulu

The creation of new terms is arguably the most efficient way whereby the scientific, educational and technical demands of a language such as Zulu can be addressed. But for terms to be created, the Zulu language has to draw upon internal and foreign sources. Among the most important methods of term creation used in Zulu that draw on internal sources aresemantic shift, derivation from verbs and idiophones and compounding. Terms drawn from foreign sources are generally known as loan words or adoptives (Wilkes and Nkosi, 1995). As with any other language pairs (for example English-Spanish or German-English), numerous problems are encountered when trying to retrieve information from different languages. The difficulty mostly lies with a 'rich' language like Zulu-that reflects its origin, history, mythology, exploits, legends, wisdom lore and world views-not all words are found in the dictionary, which makes it difficult to establish a complete word index to use. Some of the grammatical influences will now be explained below to clarify any difficulties that might exist.

3.3.4.1   Term creation by means of semantic shift

There are words that have originated because of a semantic shift. This means that the existing meaning of a word requires an expanded or modified meaning, to name a new (mostly related) concept. For instance:

Table 3.7 *Creating new terms by modifying/expanding the meaning*

Zulu word	English meaning	Modified English meaning
umnyango	door	department
isandla	hand	handwriting
bulala	kill	cancel
khala	cry	ring

One clearly sees that it would be difficult in the CLIR process to match query words to the headwords in the dictionary, if the link between the modified meaning and the original meaning is not made.

3.3.4.2   Term creation by means of compounding

Several new terms are created in compounding two or more words into a single word with a meaning that very often reflects the combined meanings of the constituted parts. For instance:

Table 3.8 *Creating new terms by combining different words*

Zulu word	English meaning	Meaning of the compound
umabonakude	television	u-ma-bona (see) + kude (far)
umakhalekhukhwini	cellular phone	u-ma-khala (cry/ring) + ekhukkwini (in the pocket)
impumalanga	east	im-phuma (come out) + ilanga (sun)
isidakamizwa	drugs	isi-daka (intoxicate) + imizwa (senses)

Compounding makes English-Zulu CLIR very difficult, because the compounded meaning is difficult to match to dictionary entries if the compounded meanings of the

different words are not linked to one another. It is not always the case that compound words are found in the dictionary itself.

3.3.4.3 Term creation by means of deideophonisation

Deideophonisation is a process that deals with the coining of terms from sounds that can be associated with the object or action that has to be named. This involves adding a prefix to the idiophone that depicts the action or the sound. For instance:

Table 3.9 *Creating new terms by adding a prefix to an associated sound*

Idiophone (without prefix)	Zulu word (with class prefix)	English meaning
bhamu (sound heard when a gun is fired)	*isibhamu*	gun
thuthuthu (sound made by motorbike)	*isithuthuthu*	motorcycle
vungu (sound made by strong wind)	*isivunguvungu*	gale wind

Because most of these terms are deduced by means of cultural interpretation, it could affect retrieval results negatively. This is mainly due to the different interpretations one would have for the different sounds made by different elements.

3.3.4.4 Term creation by means of adoptives

Many new terms in Zulu exist because they were adopted from other languages (notably from English and Afrikaans). Loanwords may differ significantly in their written form from the original words, but they sound very much the same. For example: Zulu words and English words are similar (but not identical) to each other, like *imotho–motor* (Doke et al., 1990). The loanwords may not always be listed in translation dictionaries. Many of this kind of loanwords are adapted to the phonological system of Zulu (like its syllabic structure), before they become an integral part of the Zulu lexicon. They are also known as Zululizations.

The following are a few of the many adoptives found:

Table 3.10 *Loanwords in Zulu*

Zulu word	English meaning	Zulu word	English meaning
istradi	Street derived from Afrikaans 'straat'	*imenenja*	manager
imoto	Motor/car derived from Afrikaans or English 'motor'	*ihambega*	hamburger
itafula	Table derived from Afrikaans 'tafel'	*ikhompiyutha*	computer
ithikithi	Derived from English 'ticket'	*ithemperesha*	temperature

While most of the traditional loanwords were channelled to class five (ili–) in Zulu, many of them are now channelled to class nine (in–), despite the fact that they have a class prefix connotative of class five.

Many loanwords, however, (especially place names), have been adopted with only a class prefix added to the original form. For example *iBeitbridge* and *i-Amazon*. The use of the hyphen in I-Amazon is to separate the two vowels coming together with a glottal stop between them. (See Table 5.1, Rule 8).

3.3.4.5 *Term creation by means of deverbalisation*

Although a very productive strategy of term creation in the African languages entails the process of deverbalisation, whereby nouns are derived from verbs according to certain well-known derivative rules it may have a negative impact on retrieval results. The following are a few examples:

Table 3.11 *Creating new terms by deriving verbs from nouns*

Zulu verb	English meaning	Derived Zulu word	English meaning
-*fundisa*	teach	*umfundisi*	preacher
-*zala*	give birth	*inzalo*	(financial) interest
-*thatela*	gather what has been stored by people	*intatheli*	newspaper reporter

36

3.3.4.6 Term creation by means of loan-translation

Loan-translation means that Zulu words are a literal translation of the meaning of the same word in the donor language. For instance:

Table 3.12 *Creating new terms by literally translating the word's meaning*

Zulu word	English meaning (literally)	English meaning
Indizamshini	fly machine	aeroplane
izithelo zethini	fruit of the can	canned fruit
isikhwama sokuyothenga	a bag of to go and buy	shopping bag

3.4 Dictionaries

When it comes to dictionaries, no other indigenous language in South Africa can really compare to Zulu. The number and variety of dictionaries in the Zulu language far exceed that of any other indigenous language in this country. Zulu is also the indigenous language with the oldest lexicographic history in South Africa, as it has the distinction of being the first indigenous language in which Jacob Döhne published a dictionary in 1857 (Wilkes and Nkosi, 1995). Most of the methods Döhne used (below) are still approved today:

- All the words are arranged under the root or stem letter;
- Derivatives are arranged according to their roots that show their etymology; and
- The dictionary is enriched with idiomatic sentences and phrases, illustrating the use of words.

A dictionary usually provides general translations, and it may have problems regarding aspects such as missing words, missing word forms and a lack of proper nouns (see Section 5.3 for the errors encountered with dictionaries). But the availability of bilingual dictionaries for many language pairs and the simplicity of using the dictionary for translation make it a good choice for Cross Language Information Retrieval.

3.5 *Chapter Synopsis*

This chapter was included in this study to sketch a background to an understanding of Zulu as an indigenous language, as well as to Zulu morphology. This is necessary to analyse the errors as experienced in the empirical study (described in Chapter 4) to improve retrieval performance.

This study is important to the transformation in South Africa, where language is a highly sensitive issue. It is hoped, by allowing people to use the computer in their mother tongue, it will stimulate pride in their language, as well as enabling them to learn in their mother tongue (made easier because of this). It is true that translating Zulu to English and vice versa does not remove all barriers to computer access, but it will help to eliminate some of them.

The next chapter will present a summary of the empirical work done on Zulu-English CLIR, describing the processes, the key qualitative results and findings of these results. The problems and possible solutions will be addressed in Chapter 5.