
Analysis of the Word-Initial Segment with Reference to Lemmatising Zulu Nasal Nouns

M.H. Mpungose, *Zulu Dictionary Project,*
University of Zululand, Durban-Umlazi Campus, Isipingo, South Africa

Abstract: The process of lemmatising nasal nouns in the Zulu lexicon is problematic. The traditional method is to lemmatise a Zulu lexical noun by etymological noun-stem. This practice creates difficulties in harmonising lexical nouns with their syntactic application. Most authors and dictionary-makers are inconsistent in identifying the word-initial segment which determines the letter of the alphabet under which the lexical noun should be included. Consequently, dictionary users do not find Zulu dictionaries user-friendly. This article therefore proposes the principle of "a noun without initial vowel" as a method for lemmatising Zulu nasal nouns. It concludes that it is not necessary to delve into the derivational history of a lexical noun, but rather to focus on the product of the operation of morphophonological rules. The article also suggests the need to identify the distinctiveness of the segments of a syllable and to acknowledge that identical forms of a segment do occur at different segmental positions (initial, medial and final). Finally it is argued that the Zulu nasal noun class prefix is constructed according to an open syllable pattern defined by a general CV-formula based on a VCV noun prefix open syllable pattern.

Keywords: ADJOINED LETTER, COMPOUND, COMPOSITE, CONSONANT, ELEMENT, ETYMOLOGICAL, EVOLUTIONARY, HOMORGANIC, INITIAL, INTRAVOWEL, LEMMA, LEMMATISE, LEXICAL, MORPHOPHONOLOGICAL, NASAL, NOUN CLASS PREFIX, SEGMENT, SYLLABLE, VOWEL

Opsomming: Ontleding van die woordinisiële segment met verwysing na die lemmatisering van nasale naamwoorde in Zoeloe. Die proses van lemmatisering van nasale naamwoorde in die Zoeloeleksikon is problematies. Die tradisionele metode is om leksikale selfstandige naamwoorde in Zoeloe volgens die etimologiese naamwoordstam te lemmatiseer. Hierdie gebruik veroorsaak moeïkhede by die harmonisering van leksikale selfstandige naamwoorde met hul sintaktiese toepassing. Die meeste outeurs en leksikograwe is inkonsekwent in die identifisering van die woordinisiële segment wat die letter van die alfabet bepaal waaronder die leksikale selfstandige naamwoord geplaas moet word. Gevolglik vind woordeboekgebruikers Zoeloeoordeboeke nie gebruikersvriendelik nie. In hierdie artikel word die beginsel van "n selfstandige naamwoord sonder inisiële klinker" dus voorgestel as 'n metode om nasale naamwoorde in Zoeloe te lemmatiseer. Daar word tot die gevolgtrekking gekom dat dit nie nodig is om op die afleidingsgeskiedenis van 'n leksikale naamwoord in te gaan nie, maar dat daar eerder gefokus moet word op die produk van die werking van die morfologiese reëls. Die artikel gee dit ook ter oorweging dat dit nodig is om die onderskeidende kenmerke van die segmente van 'n sillabe te identifiseer en om te erken dat identiese vorme van 'n segment in verskillende segmentele posisies (inisiël, mediaal en finaal) voorkom. Ten slotte word voorgestel dat die prefiks van die nasale

naamwoord in Zoeloe saamgestel word aan die hand van 'n oop sillabepatroon gedefinieer deur 'n algemene KV-formule gebaseer op 'n VKV-naamwoordprefiks-oopsillabepatroon.

Sleutelwoorde: NAASLIGGENDE LETTER, SAMESTELLING, SAAMGESTEL, KONSONANT, ELEMENT, ETIMOLOGIES, EVOLUSIONÊR, HOMORGANIES, INISIEEL, INTRAKLINKER, LEMMA, LEMMATISEER, LEKSIKAAL, MORFOFONOLOGIES, NASAAL, NAAMWOORDKLAS, SEGMENT, SILLABE, KLINKER

1. Introduction

Both traditional methods for lemmatising Zulu lexical items, namely the word-stem tradition and the full-word tradition, come in for criticism. These methods lead to controversies among dictionary users, dictionary makers, linguists and metalinguists. Marggraff (1997) agrees with Van Wyk's statement (1995: 82-83) that:

A dictionary is a compilation of lexical items, not a grammar ... All dictionaries therefore assume grammatical knowledge on the part of the user ... In the case of most languages a knowledge of important facts concerning the morphology and even some morphophonological processes are also regarded as a necessary prerequisite to the use of a dictionary.

The application of these two traditional methods in the Zulu dictionary-making process results in dictionaries which:

- are not user-friendly,
- assume grammatical knowledge,
- are linguistically inconsistent, and
- are uneconomical.

The basic function of a dictionary is to provide the dictionary user with comprehensible internal lexicographical information that ought to satisfy the user's needs at that moment. The lemma (head-word), as the most important dictionary entry, commands access into the required internal lexicographical information and indicates "each respective lexical unit in its canonical form" (Zgusta 1971: 249-250). The meaningful rapport between the lemma in Zulu dictionaries and Zulu dictionary users is therefore the primary concern in this analysis of the Zulu nasal noun initial segment:

Most lexicographers derive at least some satisfaction from the knowledge that the product of their labours can help ordinary language users in situations of communicative conflict or deficit (Hartmann 1983: 6).

1.1 Historical background

As early as 1857 J.L. Döhne and later also J.W. Colenso (1861, 1905), A.T. Bryant (1905, 1917), C. Roberts (1905) and R.C.A. Samuelson (1923) were already indi-

vidually confronted with the problem of lemmatising Zulu lexical items. Subsequent recent contributions, including those by Doke et al. (1948, 1958, 1964, 1990), A.C. Nkabinde (1982, 1985) and S.L. Nyembezi (1992) among others, do not offer a method different from the two traditional methods for lemmatising Zulu lexical items. The difficulty is however unfairly borne by dictionary makers who were and are expected:

- (i) to identify the Zulu lexical word as distinct from the grammatical word;
- (ii) to be conversant with the word-initial segment which determines the letter of the alphabet under which Zulu lexical items are recorded;
- (iii) to harmonise the Zulu lexicographical rules with linguistic factors;
- (iv) to coalesce lexical rules with rules that govern Zulu orthography; and
- (v) to choose a method of lemmatising words that seems best for Zulu lexical items (i.e. listing words, as Taylor (1991: 179) says, "in their appropriate phonological form").

1.2 Textual examples

In support of the hypothesis that the initial nasal consonant segment of a noun-stem is "homorganically pronounced" (Malmkjaer 1991: 26), and therefore needs not be depleted from its composite nasal consonant segment when a nasal noun is lemmatised, this article uses data taken from two Zulu monolingual dictionaries and one bilingual dictionary where more examples can be found:

- (i) Monolingual dictionaries
 - A.C. Nkabinde: *Isichazamazwi* 2 (1985) (henceforth ACNK)
 - S.L. Nyembezi: *aZ Isichazimazwi Sanamuhla Nangomuso* (1992) (henceforth SLNY)
- (ii) Bilingual dictionary
 - C.M. Doke et al.: *English – Zulu / Zulu – English Dictionary* (1990) (henceforth DMSV)

1.3 Lemmatising tendency

The three dominant tendencies adopted by most dictionary makers are:

- (i) The Zulu lemmatising practice is based on well-defined lexical theories formulated to suit languages with distinctive writing systems in which the first letter of a noun or a noun-stem is distinct. For example, Table 1 demonstrates that the initial segments in the words *pig* and *kolobe* are distinct forms. There can be no arbitrariness in lemmatising or looking up these nouns under the letter of the first consonant segment, viz. P and K respectively. In Zulu the noun *ingulube* is recorded under the letter N

by ACNK, DMSV and SLNY while Döhne (1857) lists it under the letter G. However, ACNK, DMSV and SLNY inconsistently lemmatise most nasal nouns which are classified by form with the noun *ingulube* either under one letter or under more than one letter. This tendency is evident even with the same author (see examples in Tables 9(a) and 10(a)).

Table 1: Conjunctive and disjunctive writing system

English:	The pig squealed the whole night.
Sotho:	Kolobe e tsetsetse bosiu bohle.
Zulu:	Ingulube itswininize kwaze kwasa.

- (ii) The same lexical item in Zulu is included under or excluded from the letter N or M or lemmatised under more than one letter in the case where the phonemic segment of the syllable-initial consonant of the stem, homorganically pronounced, forms the combination in the nasal compound consonant segment or the nasalised consonant segment.

Table 2: Multiletter inclusion or exclusion

The lemma for *imbuthuma* is:

- > -mbuthuma or
- > -buthuma or
- > -bhuthuma.

The lemma for *intshakaza* is:

- > -ntshakaza or
- > -tshakaza or
- > -shakaza.

- (iii) The etymological structure of the lexical item at premorphophonological level is used to determine the base-form for the canonical form of the lemma.

Table 3: The etymological lexical structure

PREMORPHOPHONO-LOGICAL LEVEL	PHONOLOGICAL PROCESS		LEXICAL ITEM
-buthuma	N + b = mb	>	<i>imbuthuma</i>
-bhuthuma	N + bh = mb	>	<i>imbuthuma</i>
-tshakaza	N + tsh' = ntsh	>	<i>intshakaza</i>
-shakaza	N + sh = ntsh	>	<i>intshakaza</i>
-hambiso	N + h = nk	>	<i>inkambiso</i>
-linganiso	N + l = nd	>	<i>indinganiso</i>

1.4 Lack of user-friendliness

These methods of lemmatising Zulu nasal nouns are inconsistent and therefore inconvenient and not user-friendly to dictionary users. The lexemes listed in the speaker's lexicon are those that are regular in his/her current language (Bauer 1988: 195). Thus the words lemmatised in a dictionary should be those that are actually used. Posthumus (1994: 35) discourages the application of etymological forms and says: "It is accepted that language does not operate with non-existing or meaningless forms" (such as those illustrated in Tables 2 and 3, e.g. **-buthuma**, **-bhuthuma**, **-tshakaza**, **-shakaza**, **-hambiso**, **-linganiso** — my examples). To use them as lemmas is therefore "unacceptable". The article postulates that lemmas need to be perceived and treated at a level of "the output of the morphological and phonological rules of the different strata put together" (Malmkjaer 1991: 323, see also Katamba 1989: 257-258). Hence the dictionary user can look up the morphophonological product but not its evolutionary process under the lemma in the dictionary.

2. A nasal consonant segment

The protoforms of "all class prefixes are bimorphic in Zulu" except class prefix 1(a) which is "truly monomorphic" (Hlongwane 1995: 62). The evolutionary noun class prefix reduction processes (Mini 1992, 1995) produce noun classes prefixed by "abstract segments" (Katamba 1989: 181, Zgusta 1971: 120) which have perceptible structures of noun class prefixes in some lexical items, for example, **-li-** in noun class (5), **-ni-** in noun classes (9) and (10), and **-lu-** in noun class (11):

noun class (5) **-li-**:

<i>indiki</i>	<i>ilindiki</i>	lemma	>	-ndiki
<i>intshontsho</i>	<i>ilintshontsho</i>	lemma	>	-ntshontsho
<i>inkankane</i>	<i>ilinkankane</i>	lemma	>	-nkankane
<i>inono</i>	<i>ilinono</i>	lemma	>	-nono
<i>inuku</i>	<i>ilinuku</i>	lemma	>	-nuku

noun class (9) **-ni-**:

<i>imbaba</i>	<i>inimbaba</i>	lemma	>	?
<i>inhlwa</i>	<i>ininhlwa</i>	lemma	>	?
<i>imvula</i>	<i>inimvula</i>	lemma	>	?
<i>intshakaza</i>	<i>inintshakaza</i>	lemma	>	?
<i>ingulube</i>	<i>iningulube</i>	lemma	>	?

noun class (10) **-ni-**:

<i>izinhlwa</i>	<i>izi-ninhlwa</i>	lemma	>	?
<i>izimvula</i>	<i>izi-nimvula</i>	lemma	>	?
<i>izintshakaza</i>	<i>izi-nintshakaza</i>	lemma	>	?
<i>izingulube</i>	<i>izi-ningulube</i>	lemma	>	?

noun class (11) -lu-:

<i>umbimbi</i>	<i>u-lumbimbi</i>	lemma	>	-mbimbi
<i>undanda</i>	<i>u-lundanda</i>	lemma	>	-ndanda
<i>unyawo</i>	<i>u-lunyawo</i>	lemma	>	-nyawo
<i>unjongwe</i>	<i>u-lunjongwe</i>	lemma	>	-njongwe
<i>umonya</i>	<i>u-lumonya</i>	lemma	>	-monya

The nasal noun classes (9) and (10) are unique and distinct. They consist of:

- the protoforms of the nasal noun class prefix which have an intravowel nasal consonant element;
- the perceptible protoforms of the nasal noun-class prefix which become homorganic with the initial consonant of the noun-stem; and
- the productive rule that governs incompatibility of adjoined consonant sounds.

The combination of protoforms of the nasal noun-class prefix with the initial consonant of noun-stems under noun classes (9) and (10) creates problems regarding the lemmatising of Zulu nasal nouns. Meinhof (1932: 95-96) states:

When a nasal is preceded by *n* < B. *ni* the *n* disappears, e.g.

n + ny > *ny*, e.g. *u-nyawo* (11) "foot" pl. *izinyawo* for **izin-nyawo*.

n + n > *n*, e.g. *izi-ne* (10) "four" > **izin-ne*.

n + m > *m*, e.g. *u-monya* (11) "python" pl. *izi-monya* for **izin-monya*, cf. *i-mini* (9) "middle of the day" pl. *izi-mini* for **izin-mini*.

The nasals remain unchanged before the vowels.

(B. denotes the protoforms of Ur-Bantu, and * denotes protoforms.)

On the basis of a proposition like *i-mini : izi-mini : *izin-mini*, this article attempts to test analogically by a reverse formula that lemmatising practice needs to ignore the morphophonological process in order to pick up initial segments of the noun-stems, that are the morphophonological products. For example, the noun *intshakaza*, grouped in the same nasal noun class (9) as *imini*, is selected here to form the problematic part of the equation, that is: If *imini : -mini*, then *intshakaza : ?*

If				
<i>imini</i> (class 9)	>	<i>i-mini</i>	:	<i>izi-mini</i> : * <i>izin-mini</i> (given) and
		<i>i-mini</i>	:	* <i>in-mini</i> (formula),
then				
<i>intshakaza</i> (class 9)	>	<i>i-ntshakaza</i>	:	<i>izi-ntshakaza</i> : * <i>izin-ntshakaza</i> and
		<i>i-ntshakaza</i>	:	* <i>in-ntshakaza</i> ;
therefore, the lemma for		<i>imini</i>	>	-mini and
for		<i>intshakaza</i>	>	-ntshakaza .

It is logical to accept that the same theory holds in noun classes (9) and (10) with all nasal nouns whose initial nasal consonant segments are either simple

or homorganic. For example, the proposed solution to find lemmata for lexical items under noun class (9) listed as examples above is:

<i>imbaba</i>	>	-mbaba
<i>inhlwa</i>	>	-nhlwa
<i>imvula</i>	>	-mvula
<i>intshakaza</i>	>	-ntshakaza
<i>ingulube</i>	>	-ngulube

Note that the evolutionary forms of the noun *intshakaza*, illustrated briefly in Table 14(a) showing developments from Ur-Bantu -Ni-, are too remote to be discussed in detail under lemmatisation. Hence the form **in-ntshakaza* is conveniently treated as example to indicate the "protoform" at premorphophonological level.

2.1 Types of nasal consonant segments

The two types of nasal consonants that occur in Zulu as consonant segments are:

- the primary nasal consonant segments, and
- the secondary nasal consonant segments.

They can be summarised schematically as in Table 4 where their relations are shown.

2.1.1 The primary nasal consonant segments

The primary or radical nasal consonant segment is found in Zulu as:

- a simple alveolar nasal consonant segment n [n];
- a simple bilabial nasal consonant segment m [m];
- a syllabic bilabial nasal consonant segment m [m̩] which occurs only if it precedes an adjoined consonant, whereas a syllabic alveolar nasal consonant is not permitted in Zulu; and
- an intravowel nasal consonant element, that is, a nasal consonant segment occurring between vowels with no adjoined consonants since a nasal consonant is never preceded by an adjoined consonant in Zulu, unless the preceding adjoined consonant is syllabic.

The features illustrated in Table 5 show no divergence in the system of lemmatisation by the Zulu dictionary makers ACNK, DMSV and SLNY. Under the primary nasal consonant segment type only the syllabic bilabial nasal consonant cluster reflects some problems. The combination of syllabic m with the initial consonant of the first syllable of the noun-stem creates indistinct forms of initial segments. This is discussed under 2.2 and shown in Tables 6, 7 and 8.

Table 4

TYPES OF NASAL CONSONANTS

RELATIONS OF THE NASAL CONSONANTS

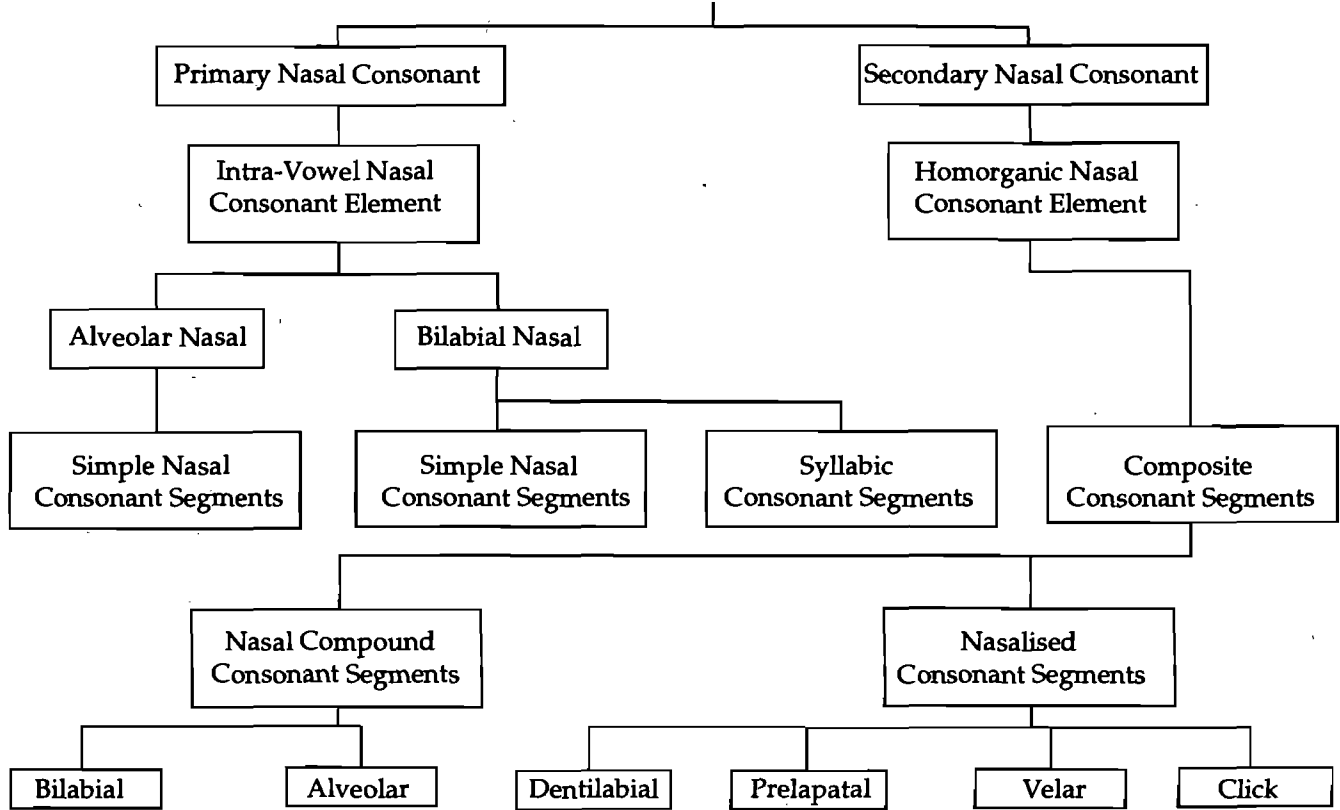


Table 5: The primary nasal consonant segments

Lexicon with Intravowel Nasal Consonant	Postulated Initial Consonant Segment	Letters under which Lemmatised by		
		ACNK	DMSV	SLNY
<i>inamfunamfu</i>	n	—	N	N
<i>inembe</i>	n	N	N	N
<i>inomfi</i>	n	—	N	N
<i>inunu</i>	n	—	N	N
<i>imamba</i>	m	—	M	M
<i>imeli</i>	m	M	M	M
<i>imenenja</i>	m	—	M	M
<i>imini</i>	m	M	M	M

2.2 Syllabic bilabial nasal consonant vs. homorganic nasal consonant segments

The ability to distinguish homographic consonant clusters between the syllabic bilabial nasal and the homorganic nasal consonant segments of written lexical items in general Zulu orthography presupposes a knowledge of the Zulu word structure, otherwise lemmatisation becomes problematic. This is attributable to current orthography. Therefore, problems in lemmatisation cannot be solved either with phonetic scripts or with specific diacritic markings for the syllabicity of segments. Hence it becomes difficult to determine or to explain the lexical rule to be applied for the initial consonant segment in both the syllabic bilabial nasal consonant and the homorganic nasal consonant segments occurring at the same segmental environment in a Zulu lexical item.

Doke et al. (1990) acknowledge the existence of this problem. For example, they divide the bilabial consonant into two dictionary alphabets, viz. **B** (or **b**) and **Ḃ**, whereby **B** is an equivalent of **Bh** (devoiced bilabial explosive) and **Ḃ** an equivalent of **B** (voiced bilabial implosive), that is, **B** = **Bh** and **Ḃ** = **B**, but **Ḃ** ≠ **Bh**, hence, **n** + **B** > **mb**, **n** + **Bh** > **mb**, and **n** + **Ḃ** > **mb** (but either **mbh** < **n** + **Bh** or **mḂ** < **n** + **Ḃ** as a single segment is incompatible). They (1990: 15, 57) assert that the two consonants **B** (or **b**) and **Ḃ** are "phonetically distinct" in Zulu:

Stems of nouns commencing in *imb* are sometimes recorded under *Ḃ*; sometimes under *mb*. When, however, it is ascertainable that the initial of the root is *b*, and in cases where the real initial is to-day unascertainable, these words are recorded under *b*.

The problem is comparatively illustrated in Table 6 by using the "minimal pair test" propounded by Katamba (1989: 22-23).

Table 6: The minimal pair test

Syllabic Nasal at Initial Consonant Clusters	Lexical Item	Homorganic Nasal Initial Consonant Segment	Lexical Item
mb	<i>imbala</i> (class 4)	mb	<i>imbala</i> (class 9)
mb	<i>umbala</i> (class 3)	mb	<i>umbalane</i> (class 3a)
mb	<i>umbethe</i> (class 3)	mb	<i>umbekle</i> (class 3a)
mf	<i>imfula</i> (class 4)	mf	<i>imfumba</i> (class 9)
mv	<i>invithi</i> (class 4)	mv	<i>invithi</i> (class 9)

Table 7: The syllabic nasal consonant segments

Lexicon with Syllabic Nasal at Initial Consonant Clusters	Postulated Initial Consonant Segment	Letters under which Lemmatised by		
		ACNK	DMSV	SLNY
<i>imbala</i>	B	—	6	B
<i>umbala</i>	B	B	6	B
<i>umbethe</i>	B	—	6	B
<i>imfula</i>	F	—	F	F
<i>invithi</i>	V	—	V	V

Table 8: The homorganic nasal consonant segments

Lexicon with Initial Homorganic Nasal Consonant Segment	Postulated Initial Consonant Segment	Letters under which Lemmatised by		
		ACNK	DMSV	SLNY
<i>imbala</i>	mb	Mb	Mb/B	Mb/Bh
<i>umbalane</i>	mb	—	Mb	Mb
<i>umbekle</i>	mb	—	B	Bh
<i>imfumba</i>	mf	—	F	Mf/F
<i>invithi</i>	mv	—	V	V

2.3 The secondary nasal consonant segments

Two secondary nasal types of consonant segments that cause a major problem in lemmatising a Zulu lexical item are:

- the nasal compound consonant segments, and
- the nasalised consonant segments.

Each type of nasal consonant segment is generated by phonological factors. Each consists of composite consonant clusters with a nasal consonant which is homorganically pronounced. Furthermore, each type of segment constitutes a single (phonetic) sound. The first actual consonant of a noun-stem therefore determines the letter under which the lexical item is recorded (see Tables 9(a) and (10)(a)). It also distinguishes each lexical item by its form and word meaning (Tables 6 and 12). The fact that the actual nasal consonant is a catalyst for nasal compound consonant segments, makes it distinct from the nasalised type of consonant segments. This feature is only realised by use of phonetic script at postmorphophonological level. The distinction is illustrated in the column "Postulated Initial Homorganic Nasal Consonant" in Tables 9 and 10.

Table 9: The nasal compound consonant segments

PREMORPHOPHONOLOGICAL LEVEL	POSTMORPHOPHONOLOGICAL LEVEL	
	Postulated Initial Homorganic Nasal Consonant Segment	Phonetic Sounds
(i) Bilabial Consonant Segments		
N + b	> mb	> [mb]
N + bh	> mb	> [mb]
N + ɓ	> mb	> [mb]
N + p	> mp	> [mp']
N + ph	> mp	> [mp']
(ii) Alveolar Consonant Segments		
N + d	> nd	> [nd]
N + l	> nd	> [nd]
N + t	> nt	> [nt']
N + th	> nt	> [nt']
N + s	> ns	> [nts']
N + z	> nz	> [ndz]
N + dl	> ndl	> [ndf]
N + hl	> nhl	> [ntf']

Table 9(a): The nasal compound consonant segments

Lexicon with Initial Homorganic Nasal Consonant Segments	Postulated Initial Homorganic Nasal Consonant Segments	Letters under which Lemmatised by		
		ACNK	DMSV	SLNY
<i>imbaba</i>	mb	Mb/-/Bh	Mb/B/Bh	Mb/B/Bh
<i>imbiza</i>	mb	Mb/-/Bh	-/-/Bh	Mb/-/Bh
<i>imbuthuma</i>	mb	-/-/-	-/-/Bh	Mb/B/Bh
<i>impandla</i>	mp	-/-/Ph	-/-/Ph	Mp/-/Ph
<i>impoko</i>	mp	-/P/-	-/P/-	Mp/-/-
<i>impontshi</i>	mp	-/-/Ph	Mp/-/-	Mp/P/Ph
<i>indaxandaxa</i>	nd	-/D/-	Nd/D/-	Nd/D/-
<i>indikimba</i>	nd	-/D/-	-/D/-	Nd/D/-
<i>indinganiso</i>	nd	-/-/-	Nd/-/L	Nd/-/-
<i>indodakazi</i>	nd	-/D/-	-/D/-	-/D/-
<i>intezazane</i>	nt	-/-/Th	Nt/-/-	Nt/-/Th
<i>intombi</i>	nt	-/-/-	-/-/Th	Nt/-/Th
<i>insada</i>	ns	-/S/-	Ns/-/-	Ns/S/-
<i>insephe</i>	ns	Ns/-/-	Ns/S/-	Ns/S/-
<i>insumpa</i>	ns	-/S/-	-/S/-	Ns/S/-
<i>inhliziyoy</i>	nhl	-/-/-	-/Hl/-	Nhl/-/-
<i>inhlw</i>	nhlw	-/Hl/-	Nhlw/Hl/-	-/Hl/-

Table 10: The nasalised consonant segments

PREMORPHOPHONOLOGICAL LEVEL	POSTMORPHOPHONOLOGICAL LEVEL	
	Postulated Initial Homorganic Nasal Consonant Segment	Phonetic Sounds
(i) Dentilabial Consonant Segments		
N + f	> mf	> [m ^h f]
N + v	> mv	> [m ^h v]
(ii) Prepalatal Consonant Segments		
N + y	> ny	> [n]
N + j	> nj	> [nd ₃]
N + sh	> ntsh	> [ntʰ]
N + tsh	> ntsh	> [ntʰ]

(iii) Velar Consonant Segments

N + g	> ng	> [ŋg]
N + k	> nk	> [ŋk']
N + kh	> nk	> [ŋk']
N + h	> nk	> [ŋk']
N + kl	> nkl	> [ŋkɫ']

(iv) Clicks Consonant Segments

N + ch	> nc	> [ŋ/]
N + c	> ngc	> [ŋ/g]
N + gc	> ngc	> [ŋ/g]
N + qh	> nq	> [ŋ!]
N + q	> ngq	> [ŋ!g]
N + gq	> ngq	> [ŋ!g]
N + xh	> nx	> [ŋ//]
N + x	> ngx	> [ŋ//g]
N + gx	> ngx	> [ŋ//g]

Table 10(a): The nasalised consonant segments

Lexicon with Initial Homorganic Nasal Consonant Segment	Postulated Initial Homorganic Nasal Consonant Segment	Letters under which Lemmatised by		
		ACNK	DMSV	SLNY
<i>imfe</i>	mf	-/F/-	-/F/-	Mf/F/-
<i>imfene</i>	mf	-/F/-	-/F/-	Mf/F/-
<i>imvakazi</i>	mv	-/V/-	-/V/-	Mv/V/-
<i>imvula</i>	mv	-/-/-	Mv/V/-	Mv/V/-
<i>inyama</i>	ny	-/-/-	Ny/-/-	Ny/-/-
<i>inyanga</i>	ny	Ny/-/-	Ny/Vowel A	Ny/-/-
<i>inyoka</i>	ny	Ny/-/-	Ny/Vowel O	Ny/Vowel O
<i>inja</i>	nj	Nj/-/-	Nj/J/-	Nj/-/-
<i>injobo</i>	nj	-/-	-/J/-	-/J/-
<i>intsha</i>	ntsh	-/-/Sh	-/-/Sh	-/-/Sh
<i>intshakaza</i>	ntsh	-/-/Sh	Ntsh/Tsh/-	Ntsh/Tsh/Sh
<i>ingaco</i>	ng	-/G/-	Ng/-/-	Ng/G/-
<i>ingobo</i>	ng	Ng/-/-	Ng/G/-	Ng/G/-
<i>inkambiso</i>	nk	-/-/H	Nk/-/H	Nk/-/H
<i>inkantolo</i>	n	-/-/-	Nx/-/Xh	Nx/-/-

3. Analysis

The dictionary entry consists of two parts (Zgusta 1971: 249-252). The first part is called the lemma (head-word), which indicates the lexical item itself. This part is most important since it acts as a lexicographic information cursor by which the dictionary entry is identified. The second part contains all information that refers to the first part. Access to the second part is never direct. It is always reached through the first part.

Bauer (1988: 9) comments on how words differ and states that:

- the *grammatical word* is discussed in terms of *its description*, and comprises the second part of the dictionary entry, e.g. verb, noun, past participle, etc., while
- the *dictionary word* is discussed in terms of *its form* (orthography or spelling) and comprises the first part of the dictionary entry, i.e. the lemma (head-word).

This article therefore analyses the first part of the dictionary entry.

3.1 The syllable and the segment

The syllable is a unit intermediate between the segment and the word. The Zulu lexical item is constructed on an open syllable pattern. For the facilitation of the lexicographic principle of lemmatisation, each syllable structure is perceived in this article to consist of a consonant segment C plus a vowel segment V, based on a CV-formula. According to Abercrombie (1980: 39, 42), a vowel is a segment of a syllable which can stand alone as a syllable to form a lone-vowel syllable, while a consonant is a segment of a syllable which, when placed alone as a syllable, becomes subjected to laws that regulate syllabic consonants and sounds collocation. This article examines the form and function of a nasal noun initial segment by which the lemma is identified in a Zulu dictionary. Abercrombie (1980: 39) asserts that "the segments of the syllable are *identified* by their sound".

3.1.1 Form and function

The algorithm is conveniently derived from Clements and Keyser (1983). It is applied with the purpose to define the form (Tables 11(a) and 11(b)) and function (Table 12) of a Zulu syllable. It forms the basis for arguments and examples throughout this article. It maintains that:

- (i) fully formed core syllables must satisfy the language-particular syllable structure conditions;
- (ii) fully formed core syllables are constructed in a pattern so that:
 - V-elements are prelinked to core syllables and C-elements to the left are adjoined one by one as long as the configuration resulting at each step satisfies all relevant syllable structure conditions;

- the syllable-initial consonants are maximised to the extent consistent with the syllable structure conditions of the language in question; and
- the construction of a syllable is in onion-like fashion, that is, built up from the centre outward (Clements and Keyser 1983: 38), hence the rule that applies to the identification of a segment can be likened to "the onion metaphor" (Katamba 1989: 258-259).

Similarly, the segment in a lexical item from the Zulu lexicon is structured in the same fashion as characterised above. The difference lies in the autonomy of the phonetic value of each segment. Each segment, either the vowel segment or the consonant segment, constitutes its own sound by which it is identifiable. The form of a syllable pattern in a lexical nasal noun can be featured as follows:

Table 11(a)	<i>intshakaza</i>	=	V + CV + CV + CV
Table 11(b)	<i>intsha</i>	=	V + CV

The first V-element represents a lone-vowel syllable and/or a vowel segment. It functions as a word-category marker (see Tables 11(a), 11(b), 14(a) and 14(b)). Hence, it influences change in lexical form and contributes to change in a word meaning. The noun in Table 11(a) is constructed of four syllables or of seven segments (4V + 3C) and in Table 11(b) of two syllables or three segments (2V + C). While proposals in this article are to:

- lemmatise a nasal noun without its initial vowel, and
- recognise the first homorganic nasal consonant in a nasal composite consonant segment of a noun-stem for lemmatisation,

the structure of a lemma can therefore for example be constructed as follows:

Table 11(a), the lemma for <i>intshakaza</i>	>	-ntshakaza (CV + CV + CV), and
Table 11(b), the lemma for <i>intsha</i>	>	-ntsha (CV).

Table 11(a): Form of a syllable and a segment

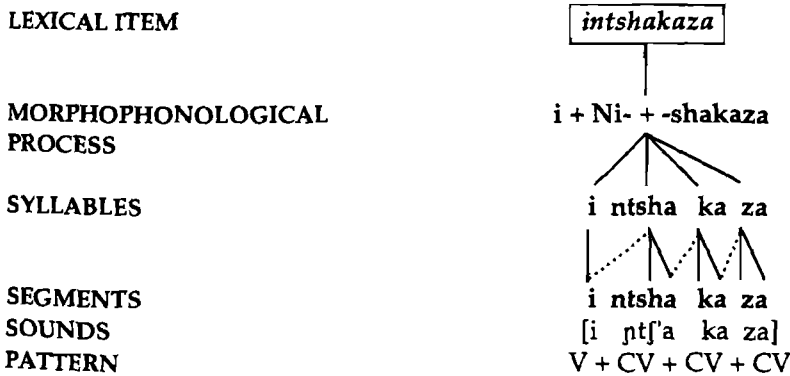


Table 11(b): Form of a syllable and a segment

LEXICAL ITEM	<div style="border: 1px solid black; padding: 2px; display: inline-block;">intsha</div>
MORPHOPHONOLOGICAL PROCESS	i + Ni- + -sha
SYLLABLES	
SEGMENTS	i ntsha
SOUNDS	[i ntʃ'a]
PATTERN	V + CV

DMSV records the word *intsha* with lemma *-sha* under the letter **S**. It lists the word *intshakaza* under the letter **N** with lemma *-ntshakaza*. ACNK and SLNY lemmatise both words *intsha* and *intshakaza* under the letter **S** with lemmas *-sha* and *-shakaza* respectively. SLNY lemmatises the word *intshakaza* under three different letters: **N** (*-ntshakaza*), **T** (*-tshakaza*) and **S** (*-shakaza*). This inconsistency is well illustrated in Table 10 (a). It is paradoxical that DMSV lists the segment *ntsh* separately as a dictionary entry under the letter **N** whereas it records the word *intsha* and many more with the same form in the initial position of the nasal noun-stem, under the letters **S** or **T**. However, Doke et al. (1996: 606) justify their inconsistency by saying:

ntsh [ntʃ] Prepalatal nasal preceding the ejective prepalatal affricate, generally the result of homorganic nasal influence upon *sh*, but sometimes upon *tsh*. (For words commencing in *intsh-* or *izintsh-* not listed under *-ntsh* see under *-sh* or *-tsh*.)

3.2 Internal cohesion of a segment

The initial nasal consonant segment in various Zulu word categories is also identifiable by form and function. It is observed in the following examples that the morphophonological product in each word category overrides the morphophonological process.

3.2.1 Syntactic application

The sentences in Table 12 illustrate a comment by Nkabinde (1992: 85) that:

The structure and function of words are closely interrelated. This makes it difficult to treat the morphology of words apart from their syntactic application.

It can be observed that the segment **ntsh** of the core syllable-initial consonant maintained its form as well as its function in each word category as illustrated in Tables 11(a), 11(b) and 12.

Table 12: Syntactic application

- (i) **Intshakaza** isiqalile ukushazwa. (Noun):
The flower-tuft has begun to wither.
- (ii) **Le ntshakaza** imhlophe qwa. (Demonstrative):
This flower-tuft is snow-white.
- (iii) **Yiphunga lentshakaza** leli olichazayo. (Possessive):
It is the flower-tuft's smell that you are describing.
- (iv) **Lo mmbila awunantshakaza.** (Negative):
This maize stalk has no flower-tuft.
- (v) **Lo mmbila usunentshakaza.** (Affirmative):
This maize plant has now grown a flower-tuft.
- (vi) **Yintshakaza** lena obuza ngayo. (Copulative):
It is the flower-tuft that you are inquiring about.
- (vii) **Izilokazane ziqaqele entshakazweni.** (Locative):
Insects adhere in a thick mass around the flower-tuft.
- (viii) **Unezinwele ezimayikayika njengentshakaza.** (Adverb):
You have hair hanging and flapping about like a flower-tuft.
- (ix) **Sengibasathe ntshakaza uyamila, lutho.** (Vocative):
I have waited for a long time for (you) the flower-tuft to grow, but there is no progress.

3.2.2 Positional mobility of a composite segment

The same nasal consonant composite segment can occupy different positions in a lexical item from the Zulu lexicon without losing its form and phonetic value. It can be placed in the initial syllable, medial syllable or final syllable positions in lexical items. The common practice of depleting a lemma's actual initial nasal consonant from its combination with the nasal consonant composite segment in its initial position in the first syllable of the noun-stem is discouraged in this article. This is demonstrated in Table 13.

Table 13: Positional mobility of a composite segment

Initial syllable:	ntsh	>	[ntʃ]
e.g. <i>intshakaza; intsha; intshibongo</i>			
Medial syllable:	ntsh	>	[ntʃ]
e.g. <i>ipentshisi; isankuntshane</i>			

Final syllable: ntsh > [ntʃ]
 e.g. *iwolintshi*; *inkwantshu*; *impontshi*

3.2.2.1 Reduplication

In the ordinary writing system, the first and second parts of reduplicated stems in a lexical item are joined to form a single lexical item which is written according to the current Zulu morphophonological form. But, when some dictionary makers lemmatise, they still change the first part of such a lexical item to reflect its form at premorphophonological level. This can be illustrated by comparing ACNK, DMSV and SLNY in Table 13(a). It can be observed that this does not greatly affect the lexical items which are categorised under primary nasal consonant segments. DMSV and SLNY demonstrate a uniform approach under this category e.g. *inamfunamfu*. Dictionary makers tend to differ when lemmatising the lexical items with duplicated stems under the secondary nasal composite consonant segments, e.g. *inkelenkele*. The postulated initial homorganic nasal composite segment for the noun *inkelenkele* is *nk*. For instance:

- ACNK records this noun under the letter K as **-KHELENKELE** and under H as **-HELE(N)HELE**;
- DMSV records it under the letter H as **-hele(n)kele** and under N as **-nke-lenkele**; and
- SLNY lists it under the letter H as **-hele(n)hele** and under N as **-nkelenkele**.

In Table 13(a), the initial segments which contrast with the postulated initial segment in each lexical item, are highlighted under each author.

Table 13(a): Reduplicated noun-stems

Lexicon with Homorganic Nasal Initial Consonant Element	Postulated Homorganic Nasal Consonant Segment	Letters under which Lemmatised by		
		ACNK	DMSV	SLNY
<i>inamfunamfu</i>	-	-	N	N
<i>imbidlimbidli</i>	mb	-/-/-	-/-/Ph	Mb/-/Bh
<i>impithimpithi</i>	mp	-/-/Ph	-/-/Ph	-/-/-
<i>indaxandaxa</i>	nd	-/D/-	Nd/D	Nd/D/-
<i>inzuthunzuthu</i>	nz	-/-/-	Nz/Z	Nz/Z/-
<i>intshipantshipa</i>	ntsh	-/-/-	Ntsh/-/Sh	Ntsh/-/Sh
<i>inkavunkavu</i>	nk	-/-/Kh	Nk/-/-	Nk/-/Kh
<i>inkelenkele</i>	nk	-/Kh/H	Nk/-/H	Nk/-/H
<i>inkinyankinya</i>	nk	-/-/Kh	Nk/-/-	Nk/-/Kh
<i>inkumunkumu</i>	nk	-/-/Kh	Nk/-/Kh	Nk/-/Kh
<i>inxakanxaka</i>	nx	-/-/-	-/X/-	Nx/X/-

3.2.3 Internal stability of a segment

The ordering of consonant items within the same nasal consonant composite segment of a syllable is fixed and noncontrastive. Hence it is not feasible to rearrange the same items and still arrive at the same or at the canonical nasal consonant composite segment. The constraint is ascribed to the rules of generative linguistics and incompatible sound laws which determine the acceptability of the morphophonological product, e.g.

- (i) N (tsh, sh) = > ntsh, but
- (ii) N (t, ts, th, ths, st, sth, ht, hs, hts) ≠ > ntsh.

Taylor (1991: 130-131), writing about the possibility of combining words into phrases, says that it is "a question of the compatibility of the feature specifications of the component forms, compatibility being in terms of selection restrictions". The acceptability of word combinations is clear: "either the feature specifications are compatible, or they are not". The same applies to the possibility of combining letters into segments, and the acceptability of these segments, in this case consonant combinations. The combination in example (ii) above does not satisfy the conditions of Zulu segment structures. The pattern of consonant configuration that results after operation in each step obeys neither the distributive law nor satisfies laws regulating the compatibility of sounds in Zulu. Therefore the operation in example (ii) cannot be acceptable in Zulu.

3.2.4 Evolutionary morphophonological product

The development of language orthography usually responds to current language demands. Logically, the current orthography of a lexical item normally indicates the first letter in the initial segment of the first syllable of the noun-stem under which the lexical item is listed in Zulu. Most dictionary-makers indicate this by a capital letter and precede it by a hyphen. When Theunissen (1943: 83) illustrates the use of capitals for the initial letter of the stem of words, he says:

It should be noted that some Class : 5 (Doke) nouns with Class : 3 (Doke) plurals, have a *different stem* in the singular than in the plural, e.g. *iNkosi* but *amaKhosi*, *iNdodakazi* but *amaDodakazi*.

The nouns *iNkosi*, *iNdodakazi* and *iNtshakaza* belong to the same noun class. This article asserts that they are, in principle, recorded under the same letter, i.e. N. In practice, however, this is not so. ACNK, DMSV and SLNY lemmatise these nouns differently. They seem to have considered the evolutionary morphophonological process more important than the lexical product itself. To prove that a lexicographical consistency exists instead of this practised inconsistency, this article provides a column for postulated initial consonant segments

in Tables 9(a), 10(a) and 13(a) that shows uniformity in the homorganic nasal composite consonant segments, viz. *nk*, *nd* and *ntsh* respectively, for these nouns.

Table 14: Evolutionary morphophonological product

(a)

	NOUN		LOCATIVE
EVOLUTIONARY	i + (Ni- + -shakaza)		e + (Ni- + -shakaza) + -ini
CURRENT FORM	i + -Ntshakaza		e + -Ntshakazeni
CONTEXTUAL FORM	<i>intshakaza</i>		<i>entshakazeni</i>
INITIAL SEGMENT	ntsh		ntsh
SOUND	[ntʃʰ]		[ntʃʰ]

(b)

	NOUN		LOCATIVE
EVOLUTIONARY	i + (Ni- + -dodakazi)		e + (Ni- + -dodakazi) + -ini
CURRENT FORM	i + -Ndodakazi		e + -Ndodakazini
CONTEXTUAL FORM	<i>indodakazi</i>		<i>endodakazini</i>
INITIAL SEGMENT	nd		nd
SOUND	[nd]		[nd]

4. Conclusion

It is difficult to engage in lemmatising in Zulu without adopting the word-stem tradition. The method for selecting the letter under which the lexical item, the nasal noun in particular, is to be recorded or looked up in the dictionary, remains controversial. Hence inconsistency, uncertainty, lack of user-friendliness, linguistic assumptions and uneconomical methods of compiling are all unavoidable. This article therefore proposes that the nasal consonant at the initial segment of the nasal noun-stem not be depleted from its homorganic composite nasal consonant segment when a nasal noun is lemmatised. All lexical items used as examples are listed in existing dictionaries, with some showing segments that are contrary to the segments shown in the column for the postulated initial homorganic nasal consonants.

Glossary of Zulu words in the text

<i>amadodakazi</i>	daughters
<i>amakhosi</i>	chiefs
<i>imbamba</i>	mamba (<i>Dendroaspis</i> spp.)
<i>imbaba</i>	the palm of the hand
<i>imbala</i> (class 4)	colours
<i>imbala</i> (class 9)	actuality, reality; fire spot on the leg
<i>imbidiimbidi</i>	heavy, stout person
<i>imbiza</i>	earthenware pot
<i>imbuthuma</i>	fire of glowing embers; large log fire
<i>imeli</i>	mare
<i>imenenja</i>	manager
<i>imfe</i>	sweet reed
<i>imfene</i>	baboon
<i>imfula</i>	rivers
<i>imfumba</i>	pile of goods
<i>imini</i>	middle of the day
<i>impandla</i>	baldness; bald-headed person
<i>impithimpithi</i>	confusion, commotion; confused affair
<i>impoko</i>	grass flower
<i>impontshi</i>	pouch; small skin bag
<i>imvakazi</i>	veil; hair-fringe hanging over the forehead (of a young woman)
<i>imvithi</i> (class 4)	species of large, shady tree
<i>imvithi</i> (class 9)	wreckage; heap of ruins
<i>imvula</i>	rain
<i>inamfunamfu</i>	sticky substance
<i>indaxandaxa</i>	person or thing dripping wet; lazy person; coward
<i>indiki</i>	person suffering from an hysterical disease
<i>indikimba</i>	subject; main facts, essential points
<i>indinganiso</i>	measure, standard for guidance
<i>indodakazi</i>	daughter
<i>inembe</i>	the last matter passed at confinement; medicine used to aid parturition
<i>ingaco</i>	cultivated field
<i>ingobo</i>	small stomach of a beast
<i>ingulube</i>	pig
<i>inhliziyo</i>	heart
<i>inhlwa</i>	flying termite
<i>inja</i>	dog
<i>injobo</i>	a strip of wild-cat's skin forming the loin-covering of a man
<i>inkambiso</i>	custom
<i>inkankane</i>	hadedda ibis (<i>Bostrychia hagedash</i>)
<i>inkantolo</i>	magistrate's court; charge office
<i>inkavunkavu</i>	course, husky kind of food

<i>inkelenkele</i>	dizziness; calamity
<i>inkinyankinya</i>	difficulty; difficult problem; difficult job; quandary
<i>inkosi</i>	chief
<i>inkumunkumu</i>	anything gritty, sandy
<i>inkwantshu</i>	cramp; contraction of muscles
<i>inomfi</i>	birdlime
<i>inono</i>	neat person
<i>insada</i>	abundance; large quantity; large number
<i>insephe</i>	springbuck
<i>insumpa</i>	wart
<i>intezazane</i>	maiden
<i>intombi</i>	girl of marriageable age
<i>intsha</i>	youth (young people)
<i>intshakaza</i>	flower-tuft
<i>intshibongo</i>	person with a long face, the lower part and frontal bone curved
<i>intshiphantshipha</i>	shy, retiring, evasive person
<i>intshontsho</i>	small piece of meat taken as a perquisite
<i>inuku</i>	untidy person
<i>inunu</i>	monster
<i>inyama</i>	meat
<i>inyanga</i>	moon
<i>inyoka</i>	snake
<i>inxakanxaka</i>	confusion, disorder; complicated structure or mechanism
<i>inzuthunzuthu</i>	supple, pliant object; calm, pleasant weather; limp, weak person or animal
<i>ipentshisi</i>	peach
<i>isankuntshane</i>	the adder-tongue fern (<i>Ophioglossum reticulatum</i>)
<i>iwolintshi</i>	an orange
<i>izimini</i>	some other times
<i>izimonya</i>	pythons
<i>izimoula</i>	rainfalls
<i>izine</i>	fours
<i>izingulube</i>	pigs
<i>izinhlwa</i>	flying termites
<i>izintshakaza</i>	flower-tufts
<i>izinyawo</i>	feet
<i>umbala</i>	colour
<i>umbalane</i>	yelloweyed canary (<i>Serinus mozambicus</i>)
<i>umbekle</i>	Cape robin (<i>Cossypha caffra</i>)
<i>umbethe</i>	dew
<i>umbimbi</i>	conspiracy
<i>umonya</i>	python
<i>undanda</i>	a tall person
<i>unjongwe</i>	acidity
<i>unyawo</i>	foot

Bibliography

- Abercrombie, D. 1980. *Elements of General Phonetics*. Edinburgh: University Press.
- Bauer, L. 1988. *Introducing Linguistic Morphology*. Edinburgh: University Press.
- Clements, G.N. and S.J. Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge: The MIT Press.
- Döhne, J.L. 1857. *Zulu-Kafir Dictionary*. Cape Town.
- Doke, C.M., D.M. Malcolm, J.M.A. Sikakana and B.W. Vilakazi. 1990. *English-Zulu/Zulu-English Dictionary*. Johannesburg: Witwatersrand University Press.
- Hartmann, R.R.K. 1983. *Lexicography: Principles and Practice*. London: Academic Press.
- Hlongwane, J.B. 1995. Growth of the Zulu Language and its Structural Changes. *South African Journal of African Languages* 15(2): 60-65.
- Katamba, F. 1989. *An Introduction to Phonology*. London: Longman.
- Lyons, J. 1971. *Introduction to Theoretical Linguistics*. Cambridge: University Press.
- Malmkjaer, K. 1991. *The Linguistics Encyclopedia*. London: Routledge.
- Marggraff, M. 1997. *Aspects of Lemmatization in Nguni*. Unpublished paper presented at the Second International Conference of the African Association for Lexicography, held at the University of Natal, Durban, 14-16 July 1997.
- Meinhof, C. 1932. *Introduction to the Phonology of the Bantu Languages*. Berlin: Dietrich Reimer.
- Mini, B.M. 1992. Problems in Lexicographical Work in the Xhosa Dictionary Project. *General and Technical Lexicography in Practice*. Pretoria: National Terminology Services.
- Mini, B.M. 1995. Lexicographical Problems in isiXhosa. *Lexikos* 5: 40-56.
- Mzolo, D. 1968. The Zulu Noun without the Initial Vowel. *African Studies* 27(4): 195-210.
- Nkabinde, A.C. 1975. *A Revision of the Word Categories in Zulu*. Unpublished D.Litt et Phil. dissertation. Pretoria: University of South Africa.
- Nkabinde, A.C. 1985. *Isichazamazwi 2*. Cape Town: Oxford University Press.
- Nkabinde, A.C. 1992. Lexicography in Zulu. *General and Technical Lexicography in Practice*. Pretoria: National Terminology Services.
- Nyembezi, S.L. 1992. *aZ Isichazamazwi Sanamuhla Nangomuso*. Pietermaritzburg: Reach Out.
- Posthumus, L.C. 1994. Word-based versus Root-based Morphology. *South African Journal of African Languages* 14(1): 28-36.
- Taylor, J.R. 1991. *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Clarendon Press.
- Theunissen, S.B. 1943. Zulu Orthography. *Native Teachers' Journal* 23: 81-83.
- Van Wyk, E.B. 1995. Linguistic Assumptions and Lexicographical Traditions in the African Languages. *Lexikos* 5: 82-96.
- Zgusta, L. 1971. *Manual of Lexicography*. Prague: Academia/The Hague: Mouton.
- Ziervogel, D. 1986. *Speech Sounds and Sound Changes of the Bantu Languages of South Africa*. Pretoria: Unisa.