

---

# Looking beyond Meaning in the *Advanced Ndebele Dictionary*\*

Langa Khumalo, *Department of Linguistics, University of KwaZulu-Natal,  
Durban, Republic of South Africa (la\_nga@yahoo.co.uk)*

---

**Abstract:** It is an established view in lexicography that the most important function of early dictionaries was to provide information on the meaning of words of a particular language. Over the years, tendencies have emerged with modern dictionaries providing detailed linguistic information resulting in more informative dictionaries. This article discusses the presentation of grammatical information, pronunciation, tone marking and usage labels and the structure and content of the back matter in the prospective *Advanced Ndebele Dictionary* (henceforth AND), which will be a successor to *Isichazamazwi SesiNdebele* (2001), the first-ever monolingual dictionary in Ndebele.<sup>1</sup> It is therefore the inclusion of this additional information that is examined in this article. The AND is still restricted to the planning stages. The work that has been done on the dictionary has been confined to academic articles about the dictionary's structure and content. The current article is a third instalment on the AND following on Khumalo 2003 and 2007.

**Keywords:** ADVANCED NDEBELE DICTIONARY, GRAMMATICAL INFORMATION, MEANING, TONE MARKING, PRONUNCIATION, USAGE LABELS, BACK MATTER

**Opsomming:** Om verder te kyk as betekenis in die *Gevorderde Ndebele-woordeboek*. Dit is 'n gevestigde siening in die leksikografie dat die belangrikste funksie van vroeëre woordeboeke was om inligting te verskaf oor die betekenis van die woorde van 'n bepaalde taal. Oor die jare het neigings na vore gekom met moderne woordeboeke wat uitvoerige taalkundige inligting verskaf wat gelei het tot meer informatiewe woordeboeke. Hierdie artikel bespreek die aanbieding van grammatiese inligting, uitspraak, toonaanduiding en gebruiksetikette en die struktuur en inhoud van die agterwerk in die *Gevorderde Ndebelewoordeboek* (voortaan GNW) wat 'n opvolger sal wees van *Isichazamazwi SesiNdebele* (2001), die heel eerste eentalige woordeboek in Ndebele. Dit is dus die insluiting van hierdie bykomende inligting wat in hierdie artikel ondersoek word. Die GNW is nog tot die beplanningstadiums beperk. Die werk wat aan die woordeboek gedoen is, het nog net gevorder tot akademiese artikels oor die woordeboek se struktuur en inhoud. Die huidige artikel is 'n derde aflewering oor die GNW in opvolging van Khumalo 2003 en 2007.

**Sleutelwoorde:** GEVORDERDE NDEBELEWOORDEBOEK, GRAMMATIESE INLIGTING, BETEKENIS, TOONAANDUIDING, UITSPRAAK, GEBRUIKSETIKETTE, AGTERWERK

---

\* This article was presented as a paper at the Fourteenth International Conference of the African Association for Lexicography, organised by the Xhosa Departement, University of the Western Cape, Bellville, Republic of South Africa, 6–8 July 2009.

## Introduction

Monolingual lexicography in Zimbabwe was initiated by the African Languages Lexical Project, known by its acronym as the ALLEX Project. The main aim of the ALLEX Project was the compilation of a variety of general and specialised monolingual dictionaries in Zimbabwe's two national languages, Shona and Ndebele. Work in the ALLEX Project, established in 1992 and institutionalised as the African Languages Research Institute in 2000, has culminated in the publication of two monolingual Shona dictionaries, *Duramazwi reChiShona* (henceforth DRC) (Chimhundu 1996) and *Duramazwi Guru reChiShona* (henceforth DGC) (Chimhundu 2001) and a general medium-sized Ndebele monolingual dictionary, *Isichazamazwi SesiNdebele* (henceforth ISN) (Hadebe 2001) (see also Chabata 2007: 281). Evidently, these dictionaries were completed in phases, with phase one being the pioneer DRC followed by phase two, which saw the production of both the DGC and the ISN. Since the ISN was preceded by the DRC, it was modelled on the latter, however, with some substantial improvements on its predecessor. Likewise the compilation of the *Advanced Ndebele Dictionary* (henceforth AND) will follow the model of its predecessor, the DGC. It is the AND, a successor to the ISN, which this article seeks to discuss.

The main objective of the ALLEX Project, which was expanded and transformed into the African Languages Research Institute (ALRI) is, according to its institutional mandate, to research, record, reanimate and develop the African languages of the country, encouraging their ever wider use in all the multifarious affairs of life by the general public of today and, increasingly, by generations to come (Fortune 2005: 1). The compilation of the AND is one of the undertakings being planned by ALRI in fulfilling this institutional mandate. As explained by Khumalo (2007: 1):

The AND will not just be larger than the ISN, but will be more advanced with regard to the depth and scope of its lexical items and definitions. ... Unlike its forerunner, the AND will provide grammatical information for each lexical item, including phonetic transcription, tone marking and etymology.

It is the inclusion of detailed grammatical information, pronunciation, tone marking and usage labels and the structure and content of the back matter in the AND that is critically examined in this article.

## The Target User

It is befitting to begin with Gouws's (2006: 85) observation that at the core of all the decisions pertaining to the compilation of any dictionary stands the person who will open the dictionary and use it, the target user. Gouws further observes that modern-day lexicography demands a clear projection of the genuine purpose of each dictionary. The fulfilment of this genuine purpose is meas-

ured by the ease with which the target user can successfully retrieve the information provided in the dictionary. It must, therefore, be clear to the lexicographer who the target user is and how best to present information to allow maximum access to knowledge and facilitate easy use by the target user.

The target groups for the ISN and AND are different. The ISN was mainly designed to fulfil the needs of lower secondary school pupils. The primary needs of the users at this level are identified as reading and comprehension, composition, basic structure of the language and general language use (Chabata and Mavhu 2005). Accordingly the ISN was crafted to contain a basic Ndebele vocabulary and a basic grammatical description with a mini-grammar in the front matter. Furthermore, it was designed to be inexpensive and easy to handle so as to be accessible to secondary school pupils. The AND on the other hand is intended to be a comprehensive reference work, which will serve as a resource for more advanced users. Its target users are those at higher levels at schools and university undergraduate level, i.e. the higher secondary and tertiary education levels. The AND is designed to be a general-purpose, medium-sized, synchronic, monolingual dictionary, meant to be a comprehensive description of the Ndebele language. Hence it is conceived to be as inclusive as possible of the vocabulary and usage of the language as gleaned from the ALLEX Ndebele corpus.<sup>2</sup> While traditionally the most central information in a dictionary is the description of words and their meaning, and that meaning is even more central when it comes to monolingual dictionaries (Landau 2001: 8), it has also become the norm in modern lexicography to include information on aspects such as the spelling, pronunciation and etymology of words. By highlighting the inclusion of this important information, this article therefore aims at looking beyond meaning in the AND.

### **Pronunciation**

A monolingual dictionary in a native language assumes that the target user's knowledge of the language is to some extent limited and/or imperfect (Svensén 1993: 20). However, it has been argued by Jackson (2002: 102) that pronunciation is not information that native speakers regularly search in a dictionary. If they do, Jackson (2002) contends, it is likely to check the pronunciation of a word that they have encountered only in writing. How a word is pronounced is very important as this is one of the idiosyncratic facts of a language.

It has been the practice of some lexicographers to proscribe the transcription of the pronunciation of what they refer to as "ordinary, everyday words". This raises the question of determining what constitutes an "ordinary, everyday word". It results in an act of pronouncing judgement on a particular aspect of the language under description. Hence the counter-argument to this has been that the exclusion of information on pronunciation represents a move away from a subservience to the recording function of general-purpose dictionaries towards considerations of what the user might or might not need (Jackson

2002: 102). Chimhundu (1997: 154) observes that "the dictionary is [...] the chief instrument in interpreting rules about spelling and pronunciation". Therefore the inclusion of pronunciation is important to a fuller description of the language.

Following Landau's (2001: 154) observation that "every lexicographer, like any good author, has his readers very much in mind", the AND will include pronunciation because the main target users would find it very helpful. The target users are tertiary and university level students who are introduced to language courses, which among others cover aspects of phonetics and phonology. Further, the dictionary may be used by second language speakers who may find the transcriptions useful to acquire the correct pronunciation of words. The inclusion of pronunciation is thus well motivated.

The transcription system to be used for representing pronunciation is the International Phonetic Alphabet (IPA). This romanized system can be used for transcribing the speech of any language. The example below shows a typical entry envisaged for the AND.

(1) **itshabi** /itʃhʌβɪ/ KKP bz 5. THUKA

The headword in example (1) is in bold lower-case roman letters which is then followed by the IPA transcription between slashes.

### Grammatical Information

It is one of the conventions of lexicography to identify the word class(es) or part(s) of speech to which each lexeme in a dictionary belongs. The traditional terms, usually abbreviated, are *ibizo* (bz) (noun), *isenzo* (sz) (verb), *isandiso* (sd) (adverb), and so on. Verbs are usually marked for their different senses such as transitive (*sz mwa*) and intransitive (*sz gmwa*). The ISN gives all this information, including number in the case of nouns. However, the AND will present more grammatical information such as tone marking for each headword, which was not provided in the ISN. It was considered that since the ISN is a monolingual dictionary, a mother-tongue speaker would hardly require the tone marking of two or more homonyms because context can easily disambiguate such words. Further, it was considered that the main target users of the ISN, i.e. secondary school pupils, are not at a level where such grammatical detail would have been introduced to them, hence complicating the accessibility of the information.

However, whereas Ndebele like most Bantu languages is tonal, the exclusion of tone marking therefore denies the complete grammatical representation of the language. It was decided therefore to include tone marking, together with the phonetic transcription of each headword in the AND, as shown in example (1). Tone, like pronunciation, is an important distinctive detail of each word in Ndebele. Example (1), here given as (2), can now be further examined.

(2) **itshabi** /itʃhəβi/ KKP bz 5. *THUKA*

The headword and its pronunciation is followed by tone marking with K, *khweza/khuphula*, for high tone, and P, *phansi* for low tone. However, as a space-saving device in the dictionary, high and low tone through-out will be marked with K followed by a hyphen (K-) for high tone and P followed by a hyphen (P-) for low tone as shown in examples (3)(a) and (b) respectively.

(3)(a) **ubaba** /uβaβa/ K- bz 1a.

(b) **mathunzi** /mat<sup>h</sup>unzi/ P- isich.

Tone marking will be followed by more grammatical information, given by an abbreviated label in standard lower-case roman letters; a number, in the case of a noun, indicating the class to which it belongs as is shown in examples (2) and (3)(a). The aim of providing grammatical information is to explain the structure of the language itself, not just the meanings of the individual words.

### Usage Labels

All dictionaries have a set of labels or style markers to indicate words or senses of words that are restricted in some way with respect to the context in which they may occur. The contextual restrictions may be geographical, i.e. depicting regional usage, as well as with regard to time (e.g. archaic), style (e.g. informal), and topic (e.g. medicine). In this section, the types of usage labels proposed for the AND are reviewed.

Dialect labels refer to geographical restrictions. The word *inopi* was particularly resisted as a legitimate entry in the ISN. It was argued that *inopi* was restricted to the Plumtree region in Zimbabwe where it was borrowed from the Kalanga language predominately spoken there. Unfortunately, because the ISN does not have usage labels, the region could not be given to indicate the restricted use of the word, although it appears in the Ndebele corpus as a loan word. It will be entered as follows in the AND.

(4) **inopi** /in<sup>h</sup>opi/ KPP bz 9. *PLUMTREE*

Plumtree is the region where this word is believed to be predominantly used. The use of the regional or dialect label would have been sufficient to assuage criticism. Related to this entry is also the inclusion in the ISN of words which were argued to be informal and should hence not be part of the dictionary. Again, the headwords in question had sufficient evidence of usage as gleaned from the corpus. But schoolteachers argued that, because the words are used only informally, they would not allow their pupils to use them in class or in their essay or composition writing. One such word in the ISN is *ishamari* which will be entered as follows in AND.

- (5) **ishamari** /iʃamari/ [itʰamali] KKPP bz 5. *KILIZA*

The label *kiliza* (informal) will indicate that the word is characteristic of conversational and casual rather than official or formal speech or writing.

The presentation of offensive words is very difficult to deal with in most cultures. A large body of offensive words in Ndebele consists of those that may be called obscene. Read (1934), quoted by Landau (1991: 229), describes obscenity as follows: "Any reference to the bodily functions that gives to anyone a certain emotional reaction, that of 'fearful thrill' in seeing, doing or speaking the forbidden. Thus, it is the existence of the ban or taboo that creates the obscenity." Offensive or taboo words refer to a variety of words that may be considered obscene, vulgar, impolite, insulting and derogatory, those referring to physical deformity or mental impairment and all those usually avoided or normally found unacceptable in the Ndebele culture. Their use goes against the norms of good behaviour because they are either injurious to one's dignity or hurt the feelings of individuals or social groups (Chabata and Mavhu 2007: 257).

Obscene words usually refer to private body parts connected with sexual activities and excretion processes and whose exposure in public is prohibited or censured. Such words are socially marked. They may not be spoken or used in public without causing embarrassment either to the speaker or the listener. Examples of offensive Ndebele words to be included in the AND are listed under (6).

- (6) *ubolo* (penis), *igolo* (vagina), *umsunu* (clitoris), *isende* (testicle), *umdididi* (anus), *-bhotsha* (defecate), *-chema* (urinate/ejaculate).

Obscene words also include terms generally considered insulting. Examples of such words are derogatory terms referring to certain ethnic groups, e.g. the Ndebele terms: *itshabi* (a foreigner or a person not of Ndebele origin), *iswina* (a Shona person) and *ihole* (a person not of Nguni origin). In Ndebele the use of offensive words in ordinary speech is prohibited and viewed as socially deviant.

Although Jackson (2002: 112) contends that there is little left in our society that is forbidden, and that some modern dictionaries no longer include usage labels, the AND will adopt the following usage labels for obscene words.

- (7)(a) **ubolo** /uβolo/ KKK bz 11. *HLONIPHISA*  
 (b) **igolo** /igolo/ KKK bz 5. *HLONIPHISA*  
 (c) **ihole** /irole/ KKP bz 5. *THUKA*  
 (d) **itshabi** /itʃaβi/ KKK bz 5. *THUKA*

Commenting on the use of such words in the first monolingual Shona dictionary, Kahari (1996–1997: 38) states that "the degree of specificity of most of the

items goes beyond what the reader who is a native speaker wants to know". The inclusion of obscene words in the dictionary, he argues, lacks "traditionally accepted taste and decorum". The AND, following the DGC, will adopt the usage labels or stylistic markers shown in example (7) (although these slightly deviate from those used in the DGC). As in previous examples, (7) presents typical entries planned for the AND.

With reference to words referring to private body parts, the AND will use the label *hloniphisa*, meaning 'embarrassing', as can be seen in examples (7)(a) and (b). Words that carry a derogatory connotation will be marked *thuka* meaning 'insulting' or 'offensive' as shown in examples (7)(c) and (d). The labels or markers will be a warning to the user that the terms are of restricted or context sensitive use. Giving such a warning is not only informative to the users but also courteous to those who might take offence.

The components of a typical major entry in the dictionary will therefore be the headword, in bold lower-case roman letters. This will be followed by the IPA transcription between slashes. Subsequently tone marking is noted, which in turn precedes further grammatical information. In case of a noun, this is rendered by an abbreviated label in standard lower-case roman letters and a number, indicating the respective noun class. The last part of the lexical entry consists of the usage label, indicating contextual restrictions such as geographical (regional), historical, and stylistic (offensive) registers. Hadebe (2007: 295) correctly observes that dictionary users usually assume and expect the dictionary to give them the 'correct' spelling, pronunciation and usage. The AND will, to a certain extent, meet some of the expectations of the users in this regard.

Modern dictionaries also often add labels indicating technical vocabulary. Svensén (1993: 49) notes that technical language arises as a consequence of constant development and specialisation in the fields of science, technology and sociology. New concepts are constantly being defined, and in order to exchange information about them, new linguistic expressions have to be coined. There are many specialised terms used in technical subjects in education, economics, sport, law, medicine and others. However, only terms that are commonly used in the Ndebele-speaking community will be candidates for selection. The prime targets are words derived from specialised fields, referred to as "international". Chitauro-Mawema (2000: 212) defines international words as those "technical words which carry specific unchanging and unambiguous senses in the contexts in which they occur and are used internationally". These are terms for respective concepts usually acquired by borrowing from other languages in which they occur. Dictionaries usually add an appropriate label to mark that a word is used in a particular field or belongs to a specific topic. The examples under (8) are typical labels that would be used in the AND.

- (8)(a) **ibhanditshi** /ibæn.ditʰi/ KPPP bz 1a *KWEZEMPILAKAHLE* (bandage – in medicine)
- (b) **ibhokiseni** /ibokiseni/ K- bz 5. *KWEZEMIDLALO* (boxing – in sport)

- (c) **ithena** /ithɛna/ KKP bz 5. *KWEZOMCULO* (tenor – in music)
- (d) **ijesi** /idʒɛsi/ KPP bz 5. *KUZEMBATHO* (jersey – clothing)
- (e) **iyabhakhasi** /ijæb<sup>h</sup>kasi/ KKPPP bz 5. *KUSAMU* (abacus – in mathematics)

By encouraging greater use of the language in various specialised fields, the dictionary is also contributing to the conducting of all kinds of education in the mother tongue. The mother tongue is the medium through which education is most effective and personalised, and through which it can best pervade society, raising levels of skill and ability with a resultant higher standard of life (Fortune 2005: 5). The AND will provide a basis on which to build specialised terminology and, while building and enriching it, be put to use in all the affairs of life from the first steps in education up to the realms of law and government, in the economy, in sport, and in the media. As Fortune (2005: 6) puts it:

The aim [...] is to see the Zimbabwean languages reach that level of development, earning and enjoying the status of a country with its own character and idiom. To do so will need a concerted commitment and effort on the part of all its speakers. But, given that, there is no reason why the African languages should not serve and deliver to their societies all that human language has done and does elsewhere.

### The Back Matter

According to Gouws (2006: 87), modern-day lexicography emphasises the need for a frame structure with front and back matter texts complementing the central list and adding venues that can accommodate lexicographic data. Users must be made aware that the dictionary as a carrier of text types presents much more than the data included in the central list. While the front matter normally carries instructions of use, describing how the dictionary is organised, how it was compiled and most importantly, how it can effectively be used, the back matter can contain practical information about language and culture that can be useful to the user's daily life.

Since the AND is a more comprehensive dictionary, its back matter will expand on the information provided in that of the ISN. For instance, the list of common first names will include information on the derivation of these names, e.g. *Musawenkosi* (God's grace) is clearly motivated by a Christian background. Further, the list of the months will also be complemented with etymological information to reveal the derivation of these names, e.g. *uNhlolanja* (literally, "dog's taboos", February) is the time for the mating of dogs, or *uNtulikazi* (literally, "a lot of dust", July) is the time characterised by much wind and dust.

Moreover, the back matter will also contain a comprehensive list of Ndebele vowels and consonants and orthographical rules. Additionally it is envisaged that it will comprise common mathematical and scientific symbols gener-



ally used in higher and tertiary institutions to introduce the language to areas where it has not been previously used. The final appendix will therefore consist of the most useful and comprehensive list applicable to the areas of linguistic analysis and literary appreciation and criticism.

## Conclusion

The article has highlighted some of the key considerations of each lexical entry in the AND, all of which go beyond just giving the meaning. The style manual of the AND states that one of the crucial information types the dictionary will provide is grammatical information, pronunciation, tone marking, usage labels and a comprehensive back matter to enhance the coverage and usefulness of the dictionary. The aim of the article has been to justify the inclusion of this information, presenting some proposed trial entries for the AND. The dictionary will hopefully encourage a wider use of the Ndebele language and by so doing also provide the basis for ultimately conducting education of all kinds in the mother tongue.

## Endnotes

1. The Ndebele language discussed here is that spoken in Zimbabwe. It is very closely related to the Zulu language. There is also the Ndebele language spoken in South Africa. There are no known studies comparing Zimbabwean and South African Ndebele.
2. The ALLEX Ndebele corpus is a systematic, well-designed and selective collection of written material and transcribed speech (Hadebe 2002: 47). The size of the corpus is currently over two million tokens (Khumalo 2007: 52).

## References

- Chabata, E.** 2007. The African Languages Research Institute: A Milestone in the Development of Zimbabwean Languages. *Lexikos* 17: 278-291.
- Chabata, E. and W. Mavhu.** 2005. To Call or Not to Call a Spade a Spade: The Dilemma of Treating 'Offensive' Terms in *Duramazwi Guru reChiShona*. *Lexikos* 15: 253-264.
- Chimhundu, H.** 1997. Language Standardisation without Policy or Planning: Zimbabwe as a Case Study. Røyneland, U. (Ed.). 1977. *Language Contact and Language Conflict*: 129-150. Volda: Volda College.
- Chimhundu, H. (Ed.).** 1996. *Duramazwi reChiShona*. Harare: College Press.
- Chimhundu, H. (Ed.).** 2001. *Duramazwi Guru reChiShona*. Harare: College Press.
- Chitauro-Mawema, M.** 2000. Grammatical Terminology for the Teaching of African Languages at Tertiary Level. *Lexikos* 10: 209-224.
- Fortune, G.** 2005. Book Review, *Duramazwi Guru reChiShona*. Unpublished.
- Gouws, R.H.** 2006. Discipline, Disciplines, Dilemmas and Data Distribution in the Planning and Compilation of Monolingual Dictionaries. *Lexikos* 16: 84-94.

- Hadebe, S.** 2002. *The Standardisation of the Ndebele Language through Dictionary Making*. Ph.D. Dissertation. Oslo: University of Oslo/Harare: University of Zimbabwe.
- Hadebe, S.** 2007. Language Planning and Monolingual Dictionaries: With Special Reference to Ndebele. *Lexikos* 17: 292-306.
- Hadebe, S. et al. (Eds.).** 2001. *Isichazamazwi SesiNdebele*. Harare: College Press.
- Jackson, H.** 2002. *Lexicography: An Introduction*. London/New York: Routledge.
- Kahari, G.P.** 1996–1997. Review of *Duramazwi reChiShona*. *Southern African Political and Economic Monthly* 10(3–4): 33-38.
- Khumalo, L.** 2003. *From a General to an Advanced Ndebele Dictionary: An Outline*. Unpublished Monograph. African Languages Research Institute.
- Khumalo, L.** 2007. *An Analysis of the Ndebele Passive Construction*. Unpublished Ph.D. Dissertation. Oslo: University of Oslo.
- Landau, S.I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. New York: Cambridge University Press.
- Read, A.W.** 1934. An Obscenity Symbol. *American Speech* 9(4): 264-278.
- Svensén, B.** 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making*. Oxford/New York: Oxford University Press.