

User-friendly Dictionaries for Zulu: An Exercise in Complexicography

Gilles-Maurice de Schryver
Ghent University

Arnett Wilkes
University of the Western Cape

In this paper the main features of Bantu lexicography are analysed through several case studies of Zulu dictionary features. Examples from both existing dictionaries as well as a forthcoming reference work are used in the analysis, which develops from verbs and nouns, gradually including more word classes, and ending with a detailed study of possessive pronouns. The latter serves as one example of the complex mappings that occur in the creation of bilingual dictionaries where the two languages involved have very different grammatical structures. In this case, one concept—that of a possessor and its possession—has only a few members in English, but hundreds in Zulu. It is shown how one can deal with such a mass of data in a structured, systematic and linguistically-sound way, all the while aiming to produce a user-friendly end product. All the members of this single concept are collectively referred to as a paradigm, and it is indicated that some members are homonymous with members of other paradigms—a fact which exponentially complicates the dictionary treatment. Several suggestions are made for the lexicographic treatment of conjunctively written Bantu languages, and all the claims, as well as all the data, are based on facts derived from a large general-language Zulu corpus.

Problem statement: User-friendly dictionaries for Zulu

Zulu, spoken in South Africa by approximately 11 million people as a home language, is one of Africa's major Bantu languages. Just like all other 500+ languages in this family, it is (a) *agglutinating* in nature, with (b) nouns assigned to different *noun classes* according to their noun class prefix, which is (c) linked to what is known as a system of *concordial agreement*. Expressed in simple terms this means that merely substituting one noun for another one from a different class will generally result in a sentence that looks (and sounds) totally different. In lexicography, this has led to a wide range of lemmatisation approaches, all of them struggling to present the reader with a user-friendly look-up method. The problem, in essence, is one of choosing the “right” morpheme(s) of each “word” for lemmatisation.

Comparing electronic dictionaries with paper dictionaries, one of course immediately realises that this problem is less acute in electronic dictionaries, as the latter can be queried in ways and directions unimaginable in a paper environment. Size (or storage space in a digital medium) also play(s) a role, where having more of it helps: cross-references (hyperlinks) can simply point the user to the entry containing the treatment. Thirdly, in dictionaries for advanced users, a detailed modular approach with massive articles several columns long can be considered.

In this paper, the hardest type is studied, namely lemmatisation in (a) a paper dictionary, (b) of modest size, (c) for young learners. In addition, (d) a bilingual Zulu – English dictionary is looked into, which adds yet another layer of complexity, as two very different grammatical systems need to be mapped onto one another. Despite these challenges, the goal is to produce user-friendly dictionaries.

Lemmatising verbs and nouns in traditional Zulu dictionaries

In order to illustrate agglutination in practice, consider *Wayesezofika ekhaya* ‘‘He was on the point of reaching home’’, which is the title of an award-winning Zulu novel (Gumbi 1966), and which can be analysed as shown in (1).¹

(1) Analysis of *Wayesezofika ekhaya* ‘‘He was on the point of reaching home’’ (Gumbi 1966)

| | | | | | | | | | | |
|-----|------|-----|--------|------|--------|--------|--------|---------|-----------|-------|
| u | a | be | e | s(e) | e | zo(ku) | fik | a | e | khaya |
| SC1 | RPTM | AUX | SC1_PM | AUX | SC1_PM | FTM | Vroot | Vending | LOCprefix | Nstem |
| | | | | | | | arrive | | | home |

u + a > wasemi-vocalisation
 be > \emptyset this auxiliary is optionally dropped in the remote past tense
 a + e > ye semi-vocalisation
 s(e) + e > se vowel elation
 zo(ku) > zo abbreviated future tense morpheme

In a so-called (traditional) ‘‘stem-based dictionary’’ users need to look up these words under the verb stem *-fika* for the verb, and the noun stem *-khaya* for the noun *ikhaya* (singular, class 5) / *amakhaya* (plural, class 6). All the other morphemes may or may not be lemmatised in the central section, and/or be brought together in the dictionary’s extra matter.

As can be seen from this straightforward example, not only is Zulu agglutinating, it is also written conjunctively—this as opposed to a disjunctive writing system, which is for instance found in Northern Sotho. For the latter, the various morphemes are physically written separately. A so-called ‘word-based dictionary’ tradition has emerged for the disjunctively written languages, where each ‘orthographic word’, even if it actually forms part of another (linguistic) word, is lemmatised. Clearly, then, lemmatising a *conjunctively* written Bantu language such as Zulu is far more difficult.

Given the multiplication potential of all the prefixes which (can) precede verbs—theoretically running into thousands of combinations for each verb—it should already be clear that the only way to realistically enter verbs into a Bantu paper dictionary, is to list them under the first letter of their stem. On this level, more or less all dictionaries agree. However, Bantu linguists who compile dictionaries will also lemmatise nouns under their noun stem, an approach which has been criticised for, amongst others, its user-unfriendliness (cf. e.g. Van Wyk 1995). If lexicographers fail to unanimously pinpoint the first letter(s) of class 9 noun stems—one of the arguments goes—how can one realistically expect dictionary users to be able to do so?² In modern, user-friendly dictionaries for the Bantu languages, nouns therefore ought to be lemmatised with (and looked up under the first letter of) their noun class (pre)prefixes.

Lemmatising other word classes in traditional Zulu dictionaries

Verbs and nouns being the main *content* words, they have by far received the most attention in the scientific lexicographic literature. One cannot use a language without the content words from other word classes, however, and certainly not without *function* words which glue all content words together. Consider, therefore, the analysis in (2) of the randomly chosen phrase *Baningi impela abahlushwa ngomamezala babo* ‘‘They are really many, the ones who are irritated by their mothers-in-law’’, lifted from the newspaper *Isolezwe* of 19 July 2004.

¹ A list of abbreviations is found at the end of the paper.

² Knowledge of Proto-Bantu (known as Ur-Bantu in South Africa) could assist lexicographers to properly identify the stem of class 9 nouns. That knowledge, however, is even more alien to the average user of a Zulu dictionary.

(2) An analysis of *Baningi impela abahlushwa ngomamezala babo* “They are really many, the ones who are irritated by their mothers-in-law” (*Isolezwe*, 19 July 2004)

| | | | | | | | | | | | |
|-----|---------|--------|-----|----------|------|---------|------------|------|---------------|-----|--------|
| ba | ningi | impela | aba | hluph | w | a | ngu | o | mamezala | ba | bo |
| SC2 | ADJstem | ADV | RC2 | Vroot | Pass | Vending | COPprefix2 | CP2b | Nstem | PC2 | Pstem2 |
| | many | really | | irritate | | | | | mother-in-law | of | their |

In a traditional stem-based dictionary, such as Doke & Vilakazi’s (1953) *Zulu – English Dictionary*, the items that can be looked up in the A to Z section are: *ba-*, *-ningi*, *impela*, *aba-*, *hlupha* [sic, without initial hyphen], —, —, *ngu-*, *o-*, *-mamezala*, *ba-*, and *-bo*. Only a user who has enough linguistic background will be able to ‘parse’ the words in this phrase, as illustrated here, so as to look up the formatives, upon which that user can then put the different meanings together so as to arrive at the overall meaning. The user of this dictionary will also need to be aware of all the possible morphophonological sound changes, here in order to for example reduce the passive verb *-hlushwa* to *-hlupha*. Further note that not all grammatical formatives have been lemmatised (nor are they listed comprehensively in the extra matter). Although Doke & Vilakazi’s dictionary remains by far the best Zulu dictionary to date, it does require advanced skills to be used. Inconsistencies in the lemmatisation approach followed for the different word classes further complicate matters a bit.

It is instructive to compare the above with the latest Zulu dictionary to reach the market, Mbatha’s (2006) *Isichazamazwi sesiZulu*, a monolingual Zulu dictionary. In that dictionary, only two words from the above phrase can be found, namely *úkúhlupha* (under the letter H; without an indication of how to form the passive however, glossing over the sound change) and *úmámezála* (under the letter M; without an indication of how to form the plural however, which is in class 2b). This outcome, clearly, is most baffling. Upon consulting the front matter of Mbatha’s dictionary, one reads the astounding claims that (a) only content words belong in a dictionary, and that (b) this means only four word classes are recognised: noun (bz), verb (sz), exclamation / interjection (bbz), and ideophone (szk), in addition to idiomatic expression (ssh) and proverb (sg). This seems to go back to Nkabinde’s PhD (1975), and is not accepted by anyone working on the Bantu languages. The results are rather catastrophic, as this latest monolingual dictionary’s only contribution to the Zulu language, then, is (the potential) that it lists a few verbs, nouns, interjections or ideophones that other Zulu dictionary makers have missed, as well as some extra idioms and proverbs.

Even more worrying is the fact that meanings were somehow forced onto extremely low-frequency to non-existing verb and noun stems. As such, one finds the noun *í(li)nîngi* (the majority) rather than the adjective stem *-ningi* (many). Likewise, the extremely-low-frequency noun *impéla* (the real one)—which is mostly used in possessive constructions, at which point it is a possessive – is found instead of the highly-frequency adverb *impela* (really). The latter has rank number 74 on a Zulu frequency list of orthographic word forms, and it is hard to argue in favour of its omission from *any* general Zulu dictionary. Consequently, and not surprisingly, a very large percentage of the Zulu language’s top-frequency words have not been lemmatised in Mbatha’s dictionary, since conjunctions, adverbs, pronouns, and so on, are not “recognised”. As an illustration, Table 1 lists the top 10 orthographic word forms in an 8.5-million-word general Zulu corpus, together with their word classes, meanings, and respective occurrences (as a percentage of all the words in the corpus).

| Rank | Word | Word class | Meaning | % |
|------|--------|--------------------------------|----------------------------------|------|
| 1 | ukuthi | conjunction | that | 1.56 |
| 2 | uma | conjunction | when; if | 0.68 |
| 3 | nje | adverb | just; simply; merely | 0.64 |
| | | adverb | in this way/manner; like this | |
| | | interjection | exactly!; just so! | |
| 4 | ngoba | conjunction | because | 0.55 |
| 5 | kodwa | conjunction | but, however, yet, nevertheless | 0.51 |
| | | exclusive quantitative pronoun | only | |
| 6 | -ke | enclitic | then, well then | 0.42 |
| 7 | lapho | adverb | there | 0.42 |
| | | conjunction | when | |
| 8 | futhi | adverb | again; once more | 0.40 |
| | | conjunction | moreover; furthermore | |
| 9 | khona | adverb | here; there | 0.38 |
| | | absolute pronoun | it (in particular); the very ... | |
| | | conjunction | so that; in order that | |
| 10 | noma | conjunction | although; even if | 0.37 |
| | | conjunction | or | |
| | | | | 5.94 |

Table 1. Top 10 orthographic word forms in Zulu (in an 8.5-million-word general Zulu corpus).

Each and every word from Table 1—together roughly six percent (5.94%) of all running words in Zulu—is missing from what is supposed to be the currently most prestigious dictionary for Zulu! The mere six-page grammatical sketch in the back matter of that dictionary certainly does not counterbalance this or all other losses in the central section.

Intermezzo: A note on the Zulu orthography

Zulu, as seen above and like all its sister Nguni languages (Xhosa, Swazi and Ndebele), has a conjunctive writing system, that is, a system whereby relatively short linguistic words are joined together to form long orthographic words with complex morphological structures. The reason why Zulu and its sister languages have adopted a conjunctive way of writing as opposed to the disjunctive method of writing followed in most other Bantu languages is not so much a scientific one as it is a practical one. Phonological processes such as vowel elision, vowel coalescence and consonant elision, all of which are mostly absent in non-Nguni Bantu languages, make the disjunctive system of writing a highly impractical one for the Nguni languages.

It is therefore absolutely necessary for Zulu lexicographers to first gain sufficient knowledge of Zulu morphology before attempting to correctly identify the boundaries of Zulu words, stems and morphemes that are to be lemmatised. Once understood, that system also needs to be described in the dictionary's extra matter, as dictionary users have to be told, more than for any other Bantu-language dictionary, how to use a dictionary for a language such as Zulu.

Lemmatising words in a novel Zulu dictionary

From what precedes it follows that, given the agglutinative morphology of Zulu combined to the conjunctive writing system, it is inevitable that one needs to make decisions regarding where to *divide* orthographic words—and thus related: how to *group* words, so canonical forms can be presented in a dictionary—in order to lemmatise the lexicon. An approach which cuts down to the smallest morpheme level (as in Doke & Vilakazi) is user-unfriendly for the target user group envisaged, while an approach which throws out most word categories, and forces so-called core Zulu meanings onto the remaining section (as in Mbatha) is even more user-unfriendly. While the former is linguistically sound, the latter moreover is not.

The user-friendly approach/solution advocated here revolves around two notions: (a) except for verbs and a few exceptions (such as the conjunction *-thi* (when), which behaves like a verb), all items from all word classes can be lemmatised *with their primary prefix(es)* included, as well as *with their suffixes* included; (b) overall *corpus frequencies* may be used in order to make a decision on the number of prefixes as well as which prefixes to include for each word class as a whole, and thus on how to organise/lemmatise the lexicon. This is indeed the approach which is followed in a new Zulu – English dictionary that is presently being finalised. Illustrated for the phrase above, each of the words/sections highlighted can directly be looked up in that dictionary: **Baningi impela abahlushwa ngomamezala babo.**

Although clearly far more user-friendly, the trade-off is that this new dictionary contains what some would view as “unnecessary repetition”. For example, instead of just one adjective stem, say *-ningi* (many; much), all the frequent full adjectival forms are included, viz. *abaningi* (class 2, many), *amaningi* (class 6, many), *eminingi* (class 4, many), *eningi* (class 9, much), *esiningi* (class 7, much), *eziningi* (classes 8 & 10, many), and *omningi* (classes 1a & 3, much). But then again, this approach enables the dictionary makers to provide carefully selected examples from the corpus for each of those forms, and thus to truly illustrate the concordial agreement system for each of the frequent classes. This approach further enables the dictionary makers to provide tailored translation equivalents for each of the forms (cf. e.g. “many” vs. “much” here).

As another example, possessive pronouns contain two formatives, written conjunctively: a possessive concord (PC) followed by a possessive stem (Pstem). See for instance *babo* in (2) above, which consists of the Pstem2 *-bo* (their, referring to ‘many’ in class 2) preceded by the PC2 *ba-* (of, in agreement with the noun “mothers-in-law”, in class 2b). The form of the *possessor* (Pstem) varies depending on the noun class or person of the object or person possessing, and likewise the form of the *possession* (PC) depends on the noun class of what is possessed. It would suffice, therefore, to list the two series of formatives, and to then assume that the user can ‘glue’ the bits together to arrive at both the correct word and its meaning for each. This is what is done in grammars, as well as in some dictionaries.

Of course, in the new Zulu – English dictionary, full forms for each possessive pronoun are presented, an approach which is surprisingly also followed by Doke & Vilakazi for this word class. Even Dent & Nyembezi (try to) do this in their *Scholar’s Zulu Dictionary* (1995). In order to get such a paradigm of related forms right, however, the argument is made in this paper that one *must* treat them all in one go. Other options, such as working through the alphabet from A to Z, or working down a lemmatised frequency list, will simply never result in a consistent treatment of each member of such a paradigm. Since an agglutinative language like Zulu contains several dozen such paradigms, some member forms of which are homonymous with one another across the paradigms, great care has to be effected in compiling each of those paradigms. Each one is complex, and given the overlaps, the multiplication of complexity can quickly get out of hand—the topic of the next section.

Lemmatising paradigms in Zulu: An exercise in complexicography

In order to illustrate this complexicography, one can briefly study the possessive pronouns in more detail. A blind application of what is found in the grammars results in all the forms shown in Table 2. As can be seen, there are (at least theoretically) $16 \times 20 = 320$ forms—all of these for what are basically just a handful of forms for the equivalent concept in English: my, your, his/her/its, our, your, and their. This is not an exceptional paradigm: one often has to ‘map’ several hundred forms from one language onto just a handful in another when bridging languages with two very different grammars in bilingual lexicography.

Corpus queries immediately indicate that some forms have a zero occurrence, and should thus not even have been mentioned in the grammars—these are the armchair words, and are printed in red (non-bold) in Table 2. Other items are simply not frequent enough to be included in a desktop dictionary – these are the rare words, and are printed in blue (bold) in Table 2. Rests all the forms highlighted in green (bold + highlight) to be treated, for which the most systematic approach to lemmatisation is to go down one column at the time. Doing this for the column headed by for instance “5” (in tandem with “11”) one can make sure that each translation equivalent is consistently indicated as “its; her/his”, or for the column headed by “1p sg” that “my” is used throughout. For the latter column, corpus data quickly reveal that each form should also be accompanied by a grammatical construction, namely “[DEM] ... ~”, meaning “of mine”—a construction which is a direct result of the need to map two different grammars.

Once one has swept through all the columns, during which one focussed on the characteristics of the possessor, one must also sweep through each of the rows, to make sure that the data relevant to each possession have been entered consistently. For PC11 (the possessive concord of class 11), for example, corpus data indicate that the same frequent combination, namely “*uqobo ~*”, meaning (depending on the possessor) “the very one(s); myself; yourself; ourselves”, should be entered throughout (except for “yourselves”, which has a zero frequency in the corpus).

| Possessor (PC) \ Possession (Pstem) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 14 | 15 | 16 | 17 | 18 | 1p sg | 1p pl | 2p sg | 2p pl |
|-------------------------------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|--------|--------|-------|--------|--------|-------|
| | khe | bo | wo | yo | lo | wo | so | zo | yo | zo | lo | bo | kho | kho | kho | kho | mi | ithu | kho | inu |
| 1 wa | wakhe | wabo | wawo | wayo | walo | wawo | waso | wazo | wayo | wazo | walo | wabo | wakho | wakho | wakho | wakho | wami | wethu | wakho | wenu |
| 2 ba | bakhe | babo | bawo | bayo | balo | bawo | baso | bazo | bayo | bazo | balo | babo | bakho | bakho | bakho | bakho | bami | bethu | bakho | benu |
| 3 wa | wakhe | wabo | wawo | wayo | walo | wawo | waso | wazo | wayo | wazo | walo | wabo | wakho | wakho | wakho | wakho | wami | wethu | wakho | wenu |
| 4 ya | yakhe | yabo | yawo | yayo | yalo | yawo | yaso | yazo | yayo | yazo | yalo | yabo | yakho | yakho | yakho | yakho | yami | yethu | yakho | yenu |
| 5 la | lakhe | labo | lawo | layo | lalo | lawo | laso | lazo | layo | lazo | lalo | labo | lakho | lakho | lakho | lakho | lami | lethu | lakho | lenu |
| 6 a | akhe | abo | awo | ayo | alo | awo | aso | azo | ayo | azo | alo | abo | akho | akho | akho | akho | ami | ethu | akho | enu |
| 7 sa | sakhe | sabo | sawo | sayo | salo | sawo | saso | sazo | sayo | sazo | salo | sabo | sakho | sakho | sakho | sakho | sami | sethu | sakho | senu |
| 8 za | zakhe | zabo | zawo | zayo | zalo | zawo | zaso | zazo | zayo | zazo | zalo | zabo | zakho | zakho | zakho | zakho | zami | zethu | zakho | zenu |
| 9 ya | yakhe | yabo | yawo | yayo | yalo | yawo | yaso | yazo | yayo | yazo | yalo | yabo | yakho | yakho | yakho | yakho | yami | yethu | yakho | yenu |
| 10 za | zakhe | zabo | zawo | zayo | zalo | zawo | zaso | zazo | zayo | zazo | zalo | zabo | zakho | zakho | zakho | zakho | zami | zethu | zakho | zenu |
| 11 lwa | lwakhe | lwabo | lwawo | lwayo | lwalo | lwawo | lwaso | lwazo | lwayo | lwazo | lwalo | lwabo | lwakho | lwakho | lwakho | lwakho | lwami | lwethu | lwakho | lwenu |
| 14 ba | bakhe | babo | bawo | bayo | balo | bawo | baso | bazo | bayo | bazo | balo | babo | bakho | bakho | bakho | bakho | bami | bethu | bakho | benu |
| 15 kwa | kwakhe | kwabo | kwawo | kwayo | kwalo | kwawo | kwaso | kwazo | kwayo | kwazo | kwalo | kwabo | kwakho | kwakho | kwakho | kwakho | kwami | kwethu | kwakho | kwenu |
| 16 kwa | kwakhe | kwabo | kwawo | kwayo | kwalo | kwawo | kwaso | kwazo | kwayo | kwazo | kwalo | kwabo | kwakho | kwakho | kwakho | kwakho | kwami | kwethu | kwakho | kwenu |
| 17 kwa | kwakhe | kwabo | kwawo | kwayo | kwalo | kwawo | kwaso | kwazo | kwayo | kwazo | kwalo | kwabo | kwakho | kwakho | kwakho | kwakho | kwami | kwethu | kwakho | kwenu |
| 18 kwa | kwakhe | kwabo | kwawo | kwayo | kwalo | kwawo | kwaso | kwazo | kwayo | kwazo | kwalo | kwabo | kwakho | kwakho | kwakho | kwakho | kwami | kwethu | kwakho | kwenu |

Table 2. Corpus frequency study of all (theoretically possible) possessive pronouns in Zulu
 As an illustration of the result of the two previous paragraphs, see the article for *lwami* shown in (3).

(3) Corpus-based treatment of *lwami* “my” in a user-friendly Zulu – English desk dictionary

lwami *possessive pronoun 11+1p sg* ► **my** Uma sengisemotweni yami noma ekhaya ngifuna ukulalela ulimi lwami. • *When I am in my car or at home I want to listen to my own language.*

▪ **[DEM] ... lwami** ► **of mine** Ngisalungiselela lolo hambo lwami. • *I'm still preparing for that journey of mine.*

▪ **uqobo lwami** ► **myself** Emehlweni azo ngibona uqobo lwami. • *In their eyes I see myself.*

As can be seen from (3), in one succinct corpus-based article, the three uses of “my”, “mine” and “myself” have been brought together.³ It needs to be remembered that this one article for *lwami* is just one cell in a two-dimensional plane, cf. Table 2. When “translated” into lexicography, that plane consists of a paradigm of about one hundred (99 to be exact) “frequent possessive-concord articles”.

As an illustration of another sweep through a row, consider the fact that the orthographic appearance of PC15 and PC16 is the same, but the meaning of class 16 possessions obscure. It was decided to add a note rather at each of the PC15 forms to indicate this. The article for *kwalo*¹, used for possessors in classes 5 and 11, and shown in (4), exemplifies this.

(4) Corpus-based treatment of *kwalo* “its; her/his” in a user-friendly Zulu – English desk dictionary

kwalo¹ *possessive pronoun 1 15+5* ► **its; her/his** “Asidingi phoyisa laseThekwini elizokhombisa ukuhlakanipha kwalo ngokuba libambe umuntu walapha eSoweto,” kusho uJonas. • *“We need no policeman here from Durban who will show us his smartness by catching a person here in Soweto,” says Jonas.* **2 15+11** ► **its; her/his** Uzocobelela ulwazi oluthile kwabanye bese uyabona ukubaluleka kwalo. • *She will seek certain knowledge from others and then realize its importance.*

Note: The pronoun ‘kwalo’ can also be used with the locative nouns of class 16, to form an adverbial phrase, in which case its meaning may become obscure.

In (4) the so-called “senses 1 and 2” treat the possessive pronouns for possessors in classes 5 and 11, with in each case a possession in class 15. The possessions in class 16 for the same possessors are covered by the Usage Note.⁴

³ Observe that the form *lwami*, being the “possessive pronoun 11+1p sg”, *has* to be used in each of the shown examples, as *ulimi* “language”, *uhambo* “journey” and *uqobo* “self” are all nouns from class 11. This, of course, is to the very core of the Bantu concordial agreement system. Also note the use of *yami*—the “possessive pronoun 9+1p sg”—in the first example under (3), which links the first person singular “my” with *imoto* “car”, a noun in class 9 (cf. also note 3). Further observe, in passing, that according to the lemmatisation strategy advocated in this paper, the form *sengisemotweni* will be look-up-able under *emotweni* ‘in the car’ rather than under the noun stem *-moto* (*i- izi-*) “car” as in traditional dictionaries.

⁴ Observe that the form *kwalo*, being the “possessive pronoun 15+5”, *has* to be used in the example under sense 1, as the possession *ukuhlakanipha* “smartness” is a noun which belongs to the infinitive class 15, while the possessor *iphoyisa* “policeman” belongs to class 5. Reformulated, and as another way to look at it, the English genitive in “the policeman’s smartness” corresponds to the Zulu possessive pronoun *kwalo*.

No doubt, consistently treating a Bantu paradigm like the one of the possessive pronouns in Zulu is no easy task, and hardly possible without professional lexicography software such as for instance *TshwaneLex*, where one can physically create the matrix from Table 2 and work through the various columns and rows in a systematic way using filters.

A comparison with the treatment of the possessive pronouns in the *Scholar's Zulu Dictionary* serves as an example. In that dictionary, the possessive pronouns are labelled '(poss)' throughout, except for all PC7 forms as well as *yenu* which are labelled "(pron)". The latter is moreover erroneously entered as *-yenu* (with hyphen). The possessive pronoun *sami* has been translated with "mine or my" (partly correct) while all other forms in the first person singular column only list "my". For *ethu* only "ours" (partly correct) is suggested, with all other forms in the first person plural column showing "our". The equivalent "hers" is listed under *zaso*, which is incorrect, as "hers" is *ezaso*. One also notices frequent alternations such as "his; her" vs. "his/her" (cf. e.g. column 1), or "its, his or her" vs. "his, her, its" (cf. e.g. *kwalo* vs. *kwaso*), and in general there is no system as to when to include his/her vs. its etc. All PC2, PC6 and PC14 forms, as well as *labo*, are missing for no apparent reason. Lastly, homonyms are sometimes mentioned, sometimes not, and when they are, the order is haphazard (cf. e.g. *kwakhe* vs. *kwabo*). There are also metalanguage typos (e.g. at *kwetho*) and inconsistencies (e.g. at *lalo*). The latter is interesting, as it is the only place where the authors added that the translation is this or that "depending on class"—obviously, that is what the mapping of this Zulu paradigm onto English is all about.

Dent & Nyembezi need to be commended for at least trying to present the possessive pronouns in a user-friendly way in their *Scholar's Zulu Dictionary*. The task is all the more daunting from the moment members of other paradigms are homonymous with the paradigm one is attempting to complete—one is easily distracted, and has to refrain from tackling several paradigms simultaneously, as some of them in turn lead to others. In this regard, possessive pronoun members are homonymous with possessives, locative adverbs, demonstrative pronouns, and an adverb.

Conclusion

Modern, user-friendly dictionaries for Zulu are within reach, as long as one has a sound grasp of the linguistic mechanisms involved, has access to corpora as well as professional lexicography software, and keeps the intended target user group in mind—at all times.

Likewise, the "possessive pronoun 15+11" *has* to be used in the example under sense 2, as the possession *ukubaluleka* "importance" is a noun which belongs to the infinitive class 15, while the possessor *ulwazi* "knowledge" belongs to class 11. Or thus, the English genitive in "the knowledge's importance" corresponds to the Zulu possessive pronoun *kwalo*.

Abbreviations used (with # the class number)

| | | | |
|---------|-----------------------|------|----------------------------|
| ADJ | adjective | Pass | passive verbal extension |
| ADV | adverb | PC# | possessive concord |
| AUX | auxiliary verb | PM | participial mood |
| COP...# | copulative | p pl | person plural |
| CP# | class prefix | p sg | person singular |
| DEM | demonstrative | RC# | relative concord |
| FTM | future tense morpheme | RPTM | remote past tense morpheme |
| LOC | locative | SC# | subject concord |
| N | noun | V | verb |
| P...# | possessive | | |

References**Dictionaries**

- Dent, G. R.; Nyembezi, C. L. S. (1995). *Scholar's Zulu Dictionary*. 3rd ed. Pietermaritzburg: Shuter & Shooter.
- Doke, C. M.; Vilakazi, B. W. (1953). *Zulu – English Dictionary*. 2nd ed. Johannesburg: Witwatersrand University Press.
- Mbatha, M. O. (2006). *Isichazamazwi sesiZulu*. Pietermaritzburg: New Dawn Publishers.

Other sources

- Gumbi, J. N. (1966). *Wayesezofika ekhaya*. S.L. [Johannesburg]: Beter Boeke. — Literary Award Winner 1966.
- Isolezwe*. (2004–2008). *Isolezwe Online Edition*. <http://www.isolezwe.co.za/> [Access date: 30 March 2008] — Zulu language newspaper.
- Nkabinde, A. C. (1975). *A revision of the word categories in Zulu*. S.L. [Pretoria]: S.N. [UNISA] — PhD Thesis.
- TshwaneLex*. (2002–2008). *TshwaneLex Suite: Professional Lexicography and Terminography Software*. <http://tshwanedje.com/tshwanelex/> [Access date: 30 March 2008]
- Van Wyk, E. B. (1995). “Linguistic Assumptions and Lexicographical Traditions in the African Languages.” *Lexikos* 5. 82-96.