Lexicographical Problems
in isiXhosa

Buyiswa M. Mini, Xhosa Dictionary Project,
University of Fort Hare, Alice

Abstract: Working on the lexicography of isiXhosa has led to the interesting experience of discovering that the description of the lexicon of isiXhosa poses innumerable challenges of both a semantic and a non-semantic nature. The problems encountered, some of which this article is an account of, include lack of understanding of the nature, depth and volume of work involved in compiling dictionaries, complexities idiosyncratic to the language itself, factors affecting consistency in methodology and other aspects, as well as extra-linguistic issues like orthography, finance, technology and relevant skills. Some of the problems rise from the very recognition of the fact that dictionaries play a significant role in "... eliminating obstacles in communication" (Alberts 1992: 1) and the consequent desire or ambition to produce as perfect a dictionary as possible.

Keywords: ISIXHOSA, TRILINGUAL DICTIONARY, MULTI-VOLUME DICTIONARY, CULTURAL CONTEXT, ADDENDA, INCLUSION, AUTHENTICITY, ACCURACY, LANGUAGE CHANGE, SELF CREATED TERMS, INDETERMINACY, ALPHABETIZATION, NATURALITY, ORTHOGRAPHY, COMPUTER LEXICOGRAPHY, RESEARCH, LANGUAGE REPRESENTATIVENESS, LANGUAGE PLANNING

Opsomming: Leksikografiese probleme in isiXhosa. Werk aan die isiXhosa-leksikografie het gelei tot die interessante ontdekking dat die beskrywing van die isiXhosa-leksikon talle uitdagings van beide semantiese en niesemantiese aard bied. Die probleme wat teengekom is, waarvan sommige in hierdie artikel beskryf word, sluit in: gebrekkige begrip vir die aard, diepte en omvang van die werk betrokke by die saamstel van woordeboeke, ingewikkeldhede eie aan die taal, faktore wat die konsekwentheid in metodologie en ander aspekte beïnvloed, asook nielinguistiese sake soos ortografie, finansies, tegnologie en relevante vaardighede. Sommige van die probleme ontstaan uit die blote besef dat woordeboeke 'n belangrike rol speel in "... eliminating obstacles in communication" (Alberts 1992: 1), en die gevolglike wens of strewe om sover moontlik 'n perfekte woordeboek daar te stel.

Sleutelwoorde: ISIXHOSA, DRIETALIGE WOORDEBOEK, MEERDELIGE WOORDEBOEK, KULTURELE KONTEKS, ADDENDA, INSLUITING, OORSPRONKLIKHEID, AKKURAATHEID, TAALVERANDERING, EIESKEPPINGS, ONBEPAALDHEID, ALFABETISERING, NATUURLIKHEID, ORTOGRAFIE, REKENAARLEKSIKOGRAFIE, NAVORSING, TAALVERTEENWOORDIGENDHEID, TAALBEPLANNING

1.   Introduction

This article highlights some of the problems encountered in the practice of lexicography in isiXhosa as well as in related issues. The main focus is on the work done at Fort Hare University's Xhosa Dictionary Project, whose three volume trilingual dictionary is called *The Greater Dictionary of Xhosa* (GDX).

One of the major problems facing lexicographers in general is the uninformed view of society that lexicography is an easy exercise, a view which may well gain support from the apparently simplified definition given by the Reader's Digest (1986: 885), though not intentionally, viz., "the writing or compilation of a dictionary or dictionaries".

This results in criticism about taking too long to finish a dictionary. Yet lexicography should be viewed, and rightly so, as part of the description of a language, which it is. Lombard (1994: 208) sees lexicography as one of the "ways of describing language".

The view stated by Fromkin and Rodman (1978: 2) (as well as others) that "in order to understand our humanity one must understand the language that makes us human" applies not only to linguistics as a science but is also directly relevant to lexicography. Viewed in this light lexicography will cease to be seen as "umdlalo wabantwana" (child's play). Taking into consideration the complexity which characterises our humanity, people will realise that lexicography too is not a simple matter of writing a word and giving its equivalent in the same or another language.

For example, the meaning of some of the entries of *The Greater Dictionary of Xhosa* is so closely tied with certain cultural aspects of the life of amaXhosa that this cultural context needs to be included in the definition of such entries as these. Sometimes it becomes necessary to give a fuller account of the cultural context than merely mentioning it, sort of in passing, in the definition. In such a case then the cultural information is catered for not in the main body of the Dictionary, but in the Addenda section (see 52 Addenda, *The Greater Dictionary of Xhosa*, Vol 3, 1989: 685-754). The cultural explanations are significant in that they help in understanding the humanity of amaXhosa by understanding their culture, contained in their language, both of which make amaXhosa human. This idea is supported and further clarified by T. Bynon who points out, though in another context, that "the lexicon is the part of a language which has the most direct links with the spiritual and material culture of its speakers and ... semantic developments may only be comprehensible by reference to the cultural background" (1977: 63).

However, the question of addenda has its own problems.

2.   Problems in connection with inclusion

A major part of lexicography practice concerns the question of which lemmas

to include and which not to, and why. The following are some of the problems encountered in this regard.

2.1   Self-created lexical items

The Xhosa Dictionary Project's first premise in making the Dictionary is that we describe the existing lexicon of isiXhosa as it occurs in the linguistic competence of amaXhosa, as evidenced mostly in their performance (cf. inter alia Langacker 1967: 34),

The implication of this premise is that we do not, or should not, create new terms even if we may feel that they need to occur in the vocabulary of isiXhosa, perhaps because concepts do exist which, it is reasonable to reckon, will soon become part of the life of amaXhosa and so will have to be referred to by means of appropriate terms of the language.

The obligation which ensues from this is that vocabulary items which may be uncommon or obscure, and therefore doubtful, are tested for their authenticity and for accuracy of the definitions we assign to them, by the competence as well as performance of the speakers of the language.

But because of personnel shortage and time limitation, one cannot claim that this happens adequately. However, these constraints aside, there would be no willingness to undertake verification activities of such terms on the part of lexicographers if such terms are self-created.

A few of such cases as referred to by Kennedy (1984: 1) as "deliberate language change for specific aims, ... to accomplish .... goals ...," do occur, however, in the GDX. Some of them have been unavoidable. In order to facilitate the lexicographical work in which we are engaged, we sometimes find ourselves faced with the necessity to create terms which have to be used in the definition of other words, objects of nature, certain concepts, or as generic terms, etc. A few examples of such originally non-existent terms, which in their turn have been made entries in the Dictionary, are the following:

(2.1.1) (a) isihluma   = something that grows; a plant, wild or not

(b) isingxaza   = liquid

(c) umthethi ⎤   = speaker/s (of a language, as distinct
                    ⎬   from isi / izithethi = people who render
(d) abathethi ⎦   speeches or arguments, say in a meeting)

(e) ithetha-siXhosa ⎤ = speaker/s of isiXhosa
                    ⎬
(f) amathetha-siXhosa ⎦

(g) isikhongozeli   = a container

(h) isikhaphi   = a form word (associated with a gram-
                               matical function)

The important question to ask in this connection is whether or not lexico-
graphers should engage in such an exercise and if so, how far should they go?
What are the guidelines and the boundaries, if any?

2.2   Forced lexical items

A related problem to the one just outlined is that of forced lexical items, which
in turn is a result of the occurrence of borrowing and of borrowed terms or
loan words (cf., inter alia, Lyons 1968: 25), with isiXhosa as the recipient lan-
guage and especially English and, to a very great extent, Afrikaans, as the
source languages. What is problematic here is not only the forced terms them-
selves but also the difficulty to determine whether some are forced or genuine
loan-words.

        The problem is further compounded by lack of exact definition of the pro-
cess of borrowing, and lack of knowledge about procedures which should be
used to identify words as loan-words and to test and prove the extent to which
loan-words have become an integral part of the recipient language. This poses
a challenge to scholars of linguistics to investigate ways in which it can be
done.

        A few examples from the work of the Xhosa Dictionary Project will illu-
strate this problem:

(2.2.1) (a) ukuprakthiza (meaning: to practise)
        (b) izonka eziyi'two' (two loaves of bread)
        (c) iposmothem
        (d) iprem
        (e) imanganiz
        (f) ilayibrari

In rural areas where isiXhosa exists in the competence of amaXhosa in its most
natural form and where bilingualism is limited to teachers and maybe the
nurses at clinics, where there are clinics, as well as to showy code-switching,
the examples in (2.2.1) are either never used or of very low frequency in the
speaker-hearers' linguistic competence and performance.

        For instance, with regards to (2.2.1)(a), if, say, pupils have sports or music
practices people say "baye emidlalweni / ezipothsini" or "baye emculweni",
respectively, the same syntactic forms they use to refer to the actual competi-
tion in sports, music, etc. (Eng.: baye = they have gone (to); emidlalweni /

ezipothsini = to the sports (practice or competition, etc.); emculweni = to music (practice or competition, etc.)). In other words, that which they have gone to practise is mentioned directly and uttered in the form of a locative adverb, instead of the verb, "- prakthiza".

(2.2.1)(c) and (e) are not known, in the cultural setting of rural amaXhosa. Words like (2.2.1)(d) and (f), if known, are usually encoded in compound nouns or whole phrases, like "into yokuqhuba umntana / usana" for (d) and "ithala leencwadi" or "indlu yeencwadi" for (f) (Eng.: (d) the thing to push a child / baby with; (e) a room / house / hut for books).

The case of (2.2.1)(b) is a rather more complicated one, which must be looked at more from a syntactic than a purely lexical point of view. The English words for the numerals, one, two, three and so on, have become an almost inherent part of the day-to-day speech (performance) of amaXhosa, not in all speech contexts, but in certain syntactic forms only. The most common is when amaXhosa refer to time. To the question by one speaker-hearer, to another "ngubani ixesha?" (what is the time?), the other will almost invariably say, in reply, "nguthu" (it is two o'clock) or "ngufo" (it is four o'clock), etc. In such cases, it is easy to see that these terms are genuine loan-terms.

However, the syntactic form presented in (2.2.1)(b) is controversial. So far as it can be ascertained this form may occur in two types of linguistic performance only. In hlonipha language, in which case it is a deliberate substitution of the English word for the one of isiXhosa which substitution is conscious and is due to social constraints rather than linguistic competence or language change. The second type is the speech of a person who has lived for long in a bilingual (or multi-lingual) setting. Again here, the substitution is quite conscious, though it may not always be deliberate, depending on the extent to which English has an influence on the speech of the individual concerned. One may further say that to say "... eziyitwo" may be an instance of style, of code-switching, especially in children.

My presentation of "two" in its English spelling and between inverted commas in (2.2.1)(b) is meant to reflect the fact that when amaXhosa substitute English words like the words for numerals, for words of isiXhosa, they are generally aware that they are mixing English with isiXhosa, perhaps for stylistic reasons or because they have temporarily forgotten the appropriate words of isiXhosa, or because there is no word of isiXhosa for a particular concept or object they wish to refer to. Because of all this, the examples in (2.2.1) seem more a case of forced loan-words than of a genuine outcome of borrowing. Mtuze (1992: 170) supports this view where he complains that "This is sometimes done so liberally that one tends to develop some skepticism about the wisdom of such wholesale borrowing."

Lombard's contention (1994: 208) that "The presentation of language in a dictionary must reflect that which is common and collective to the language of all speakers of a particular language" cuts across the acceptance of code-switching instances as dictionary entries.

Of course, a number of important questions arise here, viz.:

(i)   Should we regard cases of bilingualism as cases of genuine borrowing?
(ii)   How far is it possible to distinguish between genuine borrowing and bi-lingualism?
(iii)   Should we regard a case like that of the words for numerals as one of bi-lingualism or of borrowed terms?
(iv)   Taking into account the modern scientific understanding of the nature of language, which of the two major communities (for our present pur-poses) who speak a form of isiXhosa, viz. the rural and the urban, should we consider as more representative of natural isiXhosa?
(v)   If the concept, language, is understood to be a set of principles or rules (cf. inter alia, Smith and Wilson 1979: 13) which govern the structure or form of language in the minds of all speaker-hearers, and which thus ensure intelligibility among the members of a homogeneous group speaking the same language or language variety (cf. inter alia Langacker 1967: 23-24), do we have any justification to choose the least adulterated variety over the variety in which bi- / multi-lingual interference (see, inter alia Bynon, op cit: 239f) is most manifest, as a basis for our decision that some lexical items are proper isiXhosa and so must be entered in our Dictionary, whilst others are not and so cannot be entered? There are, possibly, many more related questions which cannot all be answered successfully yet.

2.3   Language change or not?

Tied to this problem is the question of language change which takes place in many ways, one of which is the very phenomenon of borrowing. There are words which came from English and Afrikaans which have so merged with the natural phonological and semantic subsystems of isiXhosa that we are hardly able to recognize them as loan-words. Such words as

(2.3.1) (a) itshomi (a close friend)
        (b) ukusetsha (to look for something; to investigate something)
        (c) uloliwe (train)
        (d) iswekile (sugar)
        (e) ivasi (clothes which must be or have been washed)

etc., sound so natural and their meaning so familiar that not all isiXhosa speaker-hearers can easily identify them as originating from English chum, search, railway, and from Afrikaans suiker and (om te) was, respectively.
    Because of instances like those in (2.3.1), some lexicographers and lin-guists are tempted to accept words like those in (2.2.1) as cases of language

change by the addition of lexical items. Again here, it is difficult to determine how far the idea of language change can be stretched, or what the boundaries are or should be.

Words like those in (2.2.1) may be described as mixing or language-corruption. They may or may not be the beginnings of future language change. At present they do not represent language change in so far as the lexical structure of isiXhosa is concerned. I am inclined to align myself with Langacker's view that "most new lexical items spread and come into general use much more slowly" (1967: 193), than the changes some people are only too eager to impose on isiXhosa.

On the other hand, it is perhaps wise that, because of the incredible amount of emotionalism and sentimentality which usually accompany these and related arguments, lexicographers should leave to language planners (who are usually more politicians than linguists or language-practitioners) the decision to prescribe and effect "the modification ...., the modernisation and standardisation of the lexicon" (Rubin 1984: 4).

3.   Problem of indeterminate category assignment and / or semantic content

This is a case of indeterminacy of a different kind. In this case the difficulty is in determining what syntactic categories, what parts of speech, certain words belong to, and what their meaning exactly is. So far we have not been able to come up with a solution concerning the categorization and definition of words like

(3.1)   ke (meaning: only possible in speech contexts)
(3.2)   kaloku (meaning: only possible in speech contexts)
(3.3)   akwaba (meaning: I wish that ...)
(3.4)   ngathi (eg ndibona ngathi nguye: he looks like him, etc.)
(3.5)   -anga, as in yanga iNkosi ingeva umthandazo wethu (may the Lord hear our prayer)
(3.6)   nje, as in, for example, the following utterances:
         ukuhamba nje
         ndiyabuya nje bantwana bam!
         (meaning: can only be described as a form-word, viz. isikhaphi, which term is more syntactic than semantic.)
(3.7)   qhâ, as in, for instance,
         bakhona, qha basalele (they are present, however they are still in bed)
         (meaning: however)
(3.8)   ngamana (meaning: may), e.g. ngamana iNkosi ikugcine.

Also, to define some of the words such as those mentioned under 3 above in a way that will make some sense at all we have had to resort to syntactic context and state the various meanings each of these words conveys in various sentences or phrases. For example, the only way we could define kaloku was as follows:

kàlókù
1. nakuba eli gama likholisile ukwenza umsebenzi wesihlanganisi phakathi kwesivakalisi elisingenisayo nesivakalisi, nesenzo okanye nemeko ebeseyikhankanyiwe, ikakhulu eli gama linceda ukuba ube nokuyiqonda ingqondo okanye umoya walowo uthethayo malunga nalowo athetha naye okanye naloo nto kuthethwa ngayo; ngoko ke lisetyenziswa ukubonakalisa:

1. although this word often serves as a connective between the statement it introduces and a previous statement or action or situation, it is essentially a word that serves to convey the mood or attitude of the speaker towards the person addressed or the subject under discussion; it is thus used to express:

1. hoewel hierdie woord dikwels dien as verbinding tussen 'n stelling wat dit inlei en 'n vorige stelling of handeling of situasie, dien dit eintlik om die gesindheid of luim van die spreker teenoor die aangesprokene of die saak onder bespreking aan te dui; op hierdie wyse word dit gebruik om die volgende uit te druk:

(a) imbeko: kaloku mama, andivanga ukuba uyandibiza:
(a) politeness: oh, mummy, I did not hear you call me;
(a) beleefdheid: dit spyt my, moeder, ek het u nie hoor roep nie;
(b) isicelo, isicengo: khawundincede kaloku, mhlobo wam:
(b) appeal, pleading: please help me, my friend;
(b) pleit: help my asseblief, my vriend;
(c) ugxininiso ekuchazeni okanye ekunikeni inkcazo: yimfundo kaloku eyenza ukuba sibe nempumelelo:
(c) emphasis of an explanation or of information: you see, it is education that enables us to achieve success;
(c) beklemtoning van 'n verduideliking of inligting: begryp goed, dit is opvoeding wat ons welslae laat behaal;
(d) ugxininiso ekuyaleleni: mntwana wam, kaloku kufuneka ubonise imbeko ebantwini abakhulu:
(d) emphasis of an instruction: my child, you must show respect to adults
(d) beklemtoning van 'n bevel: luister, my kind, jy moet eerbied aan alle grootmense betoon;
(e) ukukhumbuza: kaloku bawo, ubuthe uza kundithengela ilokhwe:
(e) reminding: remember, father, you said you would buy me a dress;
(e) herinnering: onthou, vader, u het gesê u sal vir my 'n nuwe rok koop;

(f) ukungqinisisa: ewe kaloku, le ndlela isinga eDikeni:

(f) affirmation: yes, of course, this road leads to Alice;

(f) bevestiging: ja, natuurlik, hierdie pad lei na Alice;

(g) ukuthundeza okanye ukukhuthaza: baleka kaloku mfondini, uyashiywa ngabanye:

(g) urging or encouraging: run, man, run; you are being left behind by the others;

(g) aanmoediging of bemoediging: hardloop, man, hardloop, jy bly agter by die ander;

(h) ukubophelela umntu kwisigqibo sakhe: ndibethe kaloku, ubuthe uya kundibetha:

(h) challenge: come on, beat me! you said you would beat me;

(h) uitdaging: slaan my dan! jy het mos gesê jy sal;

(i) ukunqanda: hayi kaloku Kholisile, musa ukusikrazula isigqubuthelo sam:

(i) admonition: no man, Kholisile, don't tear my head cloth;

(i) betraffing: nee, verbrands, Kholisile, moenie my kopdoek skeur nie;

(j) ukuthethisa: nguwe kaloku, ongcolise igama lam:

(j) reproach: it is you who have slandered my name;

(j) verwyt: dit is mos jy wat my naam belaster het;

(k) ukucaphuka, ubungxamo: arha makwenzeke leyo yenzekayo kaloku, mna ndikruqukile nguwe:

(k) annoyance, impatience: no, damn it, I am sick and tired of you now;

(k) ergernis, ongeduld: nee, genugtig, ek is nou siek en sat van jou;

(l) isisongelo: ndiya kukubetha kaloku mna:

(l) threat: you look out, I'll beat you;

(l) dreiging: ek sal jou voorwaar foeter;

(m) ukumangaliswa: ngaba kaloku yenzeka njani into yokuba nithi ninga-bemi baseJerusalem, ningazazi ezi zinto?

(m) surprise: how is it possible that you who are inhabitants of Jerusalem do not know these things?

(m) verbasing: hoe is dit moontlik dat julle wat inwoners van Jerusalem is, hierdie dinge nie weet nie?

(n) into ecacileyo okanye eyaziwayo ngulowo kuthethwa naye: sibuye singa-thenganga evenkileni; kaloku, njengokuba usazi, besingenamali:

(n) something obvious or known to the person addressed: we returned empty-handed from the shop; as you know, we had no money;

(n) vanselfsprekendheid: ons het met leë hande van die winkel teruggekeer; ons het mos nie geld gehad nie;

2. u-kaloku uyasetyenziswa njengesihlanganisi okanye igama lokubonisa apho makubekho ukunqumama ukwenzela ukuvakala kamnandi koko kuba-liswayo: ke kaloku kwakuhlwa safika mzini uthile:

2. kaloku is also used as a connective or stopword to ensure continuity and fluency in narration.

2. kaloku word ook gebruik as verbindings- of glywoord om gang en vloei-endheid in 'n vertelling te bevorder.' (*The Greater Dictionary of Xhosa*, Volume 2. Not yet published).

4.   Problems in the alphabetisation of certain nouns

Alphabetisation simply means the arrangement of lexical entries in an alpha-betical order. This, in *The Greater Dictionary of Xhosa*, is done by considering the first letter of the stem of each lexical entry. The prefix is down-played, not generally, but only in alphabetising and only for that purpose. Certain nouns of isiXhosa pose a problem in alphabetisation in that it is not easy to determine their first stem-letter. The following groups of nouns are mentioned for illus-tration:

4.1   Nouns of Classes 9 and 10

Class 9 nouns and some Class 10 nouns do not have a prefix-proper, synchroni-cally speaking. The form *-ni- which is sometimes referred to in language description is a diachronically reconstructed form (cf. inter alia Langacker 1972: 329, 352; Meinhof 1932). The N- written in capital (inter alia Welmers 1973: 165f) does not truly function as a prefix which can be separated from the stem as easily as Class 7 isi-, for example, can be separated from the stem of a noun like isikolo. Moreover, the nasal /n/ or /m/ following the pre-prefix i- in Class 9 nouns, and following the full prefix izi- and its variant ii- in nouns of Class 10, has become so homorganic with the stem-consonant which follows it that it has merged with the stem. Synchronically speaking, it is part of the stem, at least in isiXhosa. Also, if we consider the little theory about the evolution of nasal compounds in languages like isiXhosa, which theory is ascribed to Carl Meinhof, it is logical to assume that nouns like inkomo, impempe, etc. (in Class 9) and izintso, iintambo, iimpumlo, etc. (in Class 10) all have the nasal /n/ or /m/ as the first letter of the stem. Therefore, in the Xhosa Dictionary Project we find it more natural to alphabetise such nouns under the letters N or M rather than under those letters that follow N or M. The same applies to deverbative nouns of these classes as well as plural forms (which are in Class 10) of nouns which do not have a nasal as the first letter of their stem in the singular: e.g.

deverbatives: intsalela (remnant) < ukusalela
                imfuyo (livestock) < ukufuya

plural forms: iintsana (babies) < usana
                 iimpawu (signs, characteristics) < uphawu.

However, my predecessor felt that they must be under, say k, p, t, f, etc. as the case may be. But since this is not in harmony with the intuitions of many isiXhosa speaker-hearers either, be they unschooled or students of the language, including mine, we have decided to alphabetise these nouns under either the letter N or the letter M. Consequently there are some nouns of Classes 9 and 10 which at present do appear in Volume 3, the only volume of our Dictionary published yet, but which will appear in Volume 2 as well, since Volume 2 will contain the letters K to P.

4.2   Nouns the stems of which appear to commence with vowels

It is difficult to identify the first letter of the stem in nouns like

(i)   Ulwandle (sea)
(ii)   ulwamvila (sting)
(iii)   ulwabhici (confusion; intricate / difficult matter)
(iv)   isandla (hand)
(v)   amehlo (eyes)

This difficulty thwarts attempts at simplicity, explicitness and user-friendliness which are some of the main guiding principles in the making of *The Greater Dictionary of Xhosa*. And the only solution we have come to, so far, which is far from being the best, has been recourse to technicality, in direct contravention of the criterion that "the less technical the linguistic knowledge presupposed by a dictionary, the more user-friendly it will be" (Van Wyk 1989, in Mtintsilana, 1990: 9).

    We have had to apply a morphological analysis to come to the conclusion that in the case of (i), (ii), (iii) and (iv) above, the first stem-letter is A. The case of (v) was a bit tricky in that with a genuine application of morphological analysis we ended up with I as the initial letter of the stem of the noun amehlo (cf. ama + (i)hlo).

    Then this sounded most unnatural and so we adopted, quite arbitrarily, the attitude that for the purposes of the Dictionary only, the first letter of the stem of this noun is H.

5.   Other problems relating to form

5.1   Adjectives and relatives

The problems elucidated in 4 above are related to the form or structure of the different entries are meant to assume in the Dictionary. A similar problem is encountered with regards to adjectives and relatives, this time not with the shape or alphabetical position of these lexical items themselves, but with the form or structure of some of the key words used in defining them. My argument here is based on examples like the following:

(5.1.1) '-nyhidi-nyhidi bj / cop; -nyhìdì-nyhìdì bl / rel:

1.   -dyojwe amafutha yachininika; okwamehlo okanye ubuso: -manzi ziinyembezi; okomzimba: -manzi dyubhu kukubila:
1.   smeared with grease or fat; of the eyes or face: soiled or wet with tears; of the body: wet with perspiration;
1.   besmeer met vet of teer; van die oë of gesig: vuil of traanbesmeerd; van die liggaam: natgesweet;
2.   okomntu: -nesiqu esikhulu ekhohlwe kukuhamba:
2.   of a person: corpulent, obese, very fat.
2.   van 'n persoon: uitermate vet, swaarlywig, drillend van vet. (*The Greater Dictionary of Xhosa*, Vol 2: not yet published).

(5.1.2) '-rholofu bj / cop b1 / rel < Afr:

1 okoboya begusha: -ngathambile, -ngasongananga, -ntsaka-ntsaka' ..."
(Pahl, et al, 1989: 130).

The forms in question are the verbs used without their concords, viz. -dyojwe ... in (5.1.1) and -ngathambile and -ngasongananga in (5.1.2) as well as the shortened relative copulative, -nesiqu ... in example (5.1.1).

These forms represent incomplete utterances, and are cut in syllables where a natural speaker of the language would not normally shorten them. The sense they convey is reduced, thus sacrificing authenticity of form, so far as the language is concerned, for technicality. It being our editorial policy to present the language as naturally as possible in the definitions, thus to maintain and preserve naturality, a change of form was thus proposed, though when presented to interested outside people, it was not received with any zeal and yet not rejected outrightly. This change involves the use in the definition of the qualifiers in question, of full nouns in some cases and of infinitive forms in many other cases. Thus the qualifier in example (5.1.1) would be defined in the following way:

-nyhìdi-nyhìdì bl / rel, bj / cop: isibaluli, isibanjalo / isichazi esichaza:
1.   ukuchininika kwento edyojwe amafutha kakhulu; ukuba manzi kobuso
       okanye amehlo ziinyembezi; ukuba manzi dyubhu komzimba kukubila:
2.   ukuba nesiqu esikhulu, de ukhohlwe kukuhamba:

And the definition of the qualifier in (5.1.2) would have a format like the fol-
lowing:

-rholofu bl / rel, bj / cop < Af: 1 isichazi / isibaluli, isibanjalo esichaza:
ukungathambi, ukungasongani, ukuba ntsaka-ntsaka kwento enjengo-
boya begusha; etc.

As has been said, this method is not readily welcomed. It is felt, especially by
those who have for a long time been familiar with the form I am inclined to
reject, that with the use of infinitive forms there is little distinction now
between the definition of verbs and that of relatives and adjectives. However, I
feel that there is a difference, and that however marginal, it is significant in dis-
tinguishing between verbs and qualifiers as entries in the Dictionary.

6.   Orthography-related problems

6.1   The orthography of compounds

The general public's attitude toward a dictionary as a source of "accurate" lan-
guage and thus as an influencing factor in the standardization of a language,
the fact that a dictionary has an influence on people's use of language (Mtuze
1992: 166), as well as the fact that lexicographers of isiXhosa see their work as
preservation of isiXhosa as a language, all place a heavy responsibility on the
lexicographers to ensure that they maintain a very high standard of accuracy in
all aspects of their work. One of the obligations which result from this outlook
is that of dealing with some inconsistencies in orthography.   The most
problematic case in this regard is the presentation of compounds.
       The superficial treatment of the orthography of compounds in the "blue
book", the *Xhosa Terminology and Orthography No. 3* of the then Department of
Bantu Education (1972: 35, 37) gives the impression that it is easy to simply
write compounds with an apostrophe or a hyphen, with the option to leave out
these writing devices at will. This option, which we find exercised by authors
and other users or practitioners of isiXhosa, results in inconsistency in the
presentation of compound words as lemmas in the dictionary.
       Besides the problem of inconsistency, we have another consideration, viz.
that of enabling the readers / users of our dictionary to identify the two or
more lemmas which combine to form each compound. The need for the easy
identification becomes more urgent when we consider the trilingual nature of

our dictionary and thus its potential contribution to communication in the multi-lingual country that South Africa is. In some kinds of compounds because of their nature, this identity is clear, e.g.

> ubulawu bamagqirha
> ubulawu obumhlophe / ubulaw' obumhlophe

In others it is not so clear, e.g.

> isijamankungwini
> ungxowayizali
> *idlakudla*, etc.

Not only because of the knowledge that a dictionary is not only a semantic description of language but also a guide to other aspects of language study, and of the influence of the public's view — as well as that of ours — that isiXhosa lexicographers are among the most influential and important custodians of the language, but also for consistence in our own work, we have decided to present compounds like the latter three above either with a hyphen, or an apostrophe, whichever is appropriate, as in

> isijama-nkungwini and idla-kudla, and ungxow' ayizali, respectively.

But the problem does not end there. There are compounds which it has become a norm to write conjunctively, and this happens predominantly in the written literature of isiXhosa. Cases like

> *ndawonye,* ngandletyananye, untlalontle untsukumbini, *oonyawontle,* *ndlelantle,* unamnawe, etc.,

illustrate this position. The "blue book" says that they are written "conjunctively or hyphenated" and that "hyphenation is optional" (Department of Bantu Education 1972: 35).

Even to us, it seems natural to write some of them conjunctively, thus treating them as consisting of a single lexeme, as in the case of those which are italicized in the paragraphs above. I do not know if their appearance, in Volume 2 or in Volume 1 in the following hyphenated form, will not raise a storm from both professional and lay-users of the dictionary:

> idla — kudla
> ndawo — nye
> unyawo — ntle
> ndlela — ntle, etc.

6.2   Where to put capital letters

It is a stated editorial policy of the GDX that lemmas are entered in appropriate alphabetic positions by considering the first letter of the stem. The exception is where a word has become a proper noun as in

> uLizwi (Jesus Christ), uMlungisi (a male person's name), uMongameli (a male person's name or title).

Treated as ordinary lemmas the above cases begin with Z, L and O respectively. However, to revere God, all terms referring to Him or to something connected with Him are written with capitals even in ordinary discourse. Here is an interesting example:

> iLizwi likaThixo laphefumlelwa nguYe. (The Word of God was inspired by Him)

ILizwi and nguYe are ordinary words and not proper nouns. It is my contention that the placement of the capital in "nguYe" is appropriate, but not in iLizwi, where it should be on the stem-initial consonant Z, and not on the prefixal element, viz. L.

To me, the same applies to titles given to books, to people holding certain positions, etc. Thus, cases like

> abaPostile, amaKrestu, abaFundisi, abaHloli, abaNumzana, ubuRhulumente, ubuKumkani, isiKolo saseLovedale, abaPhathiswa. (Department of Bantu Education 1972: 27).

are appropriate, whereas the following cases

> iMbongi yeSizwe Jikelele, uMfundisi, uMlungiseleli, uMphathiswa weze-Mfundo, iZenzo zabaPostile, uMongameli weMfundo, aBongameli beMfundo, iLizwi likaThixo; iSitya seSizwe (Ibid).

would have been appropriate, and will be entered in the GDX, with the capitals on the first letter of the stem, as follows

> iMbongi yesiZwe Jikelele, umFundisi, umLungiseleli, umPhathiswa, umPhathiswa wezeMfundo, izEnzo zabaPostile, umOngameli weMfundo, abOngameli beMfundo, iliZwi likaThixo, isiTya sesiZwe, etc.

7.   Automation of data

As pointed out by Mtintsilana (1990: 20 et al) there is very little that has been done in our Dictionary Project in this regard. Initially, isolation and therefore lack of knowledge concerning computational lexicography were the chief culprits. At present we are on our way to building an electronic data-base. And the problems we face in this regard now include lack of trained personnel to do this. This, in turn, is delaying the progress of the work as time has to be spent on the empowerment of, especially, the editorial staff. On the positive side, our linking to the University's local area network has ultimately become a reality. And the generous assistance offered by the Bureau of the Woordeboek van die Afrikaanse Taal, of Stellenbosch, since the beginning of this year (1995) is appreciated.

8.   Problems with research

Apart from the familiar problems associated with field research, which is the backbone of our Dictionary work, shortage of staff to engage fully in research and shortage of funds to get more staff are some of the problems in this connection. At present there are only two people responsible for field- and other research, in addition to their heavy load of editorial duties. The foremost hurdle, however, is again lack of funds to provide for transport to all the isiXhosa speaking areas, and for other expenses directly connected with the research. This situation is to be lamented because research, especially field research, is the most important way in which we can ensure that our work is authentic and is on the right track, at least so far as representativeness of all varieties of isiXhosa is concerned.

9.   Conclusion

Various problems relating to orthography in isiXhosa as experienced by the lexicographers of the Xhosa Dictionary Project of the University of Fort Hare (situated in Alice, in the Eastern Cape) have been highlighted albeit inexhaustively. There are other concerns of equal dimension which have not been included here because of the limited scope of this paper. Some of the problems discussed have implications for, or solutions in, language planning; others have implications for orthography. Some pertain to editorial policy (e.g. inclusion), some emanate from the structure of the language itself, whilst others are extra-linguistic.

References

Alberts, Mariëtta. 1992. Behoeftebepaling in die leksikografie. *Lexikos* 2 (AFRILEX-series 2: 1992): 1-27.

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

Department of Bantu Education. 1972. *Xhosa Terminology and Orthography No. 3.* Pretoria: Government.

Fromkin, V. and Rodman, R. 1978. *An Introduction to Language.* 3rd ed. New York: Holt, Rinehart and Winston.

Kennedy, C. (Ed). 1984. *Language Planning and Language Education.* London: George Allen & Urwin.

Langacker, R.W. 1967. *Language and Its Structure.* 2nd ed. New York: Harcourt Brace Jovanovich.

Langacker, R.W. 1972. *Fundamentals of Linguistic Analysis.* New York: Harcourt Brace Jovanovich

Lombard, F.J. 1994. Lexicographer, Linguist, Dictionary User: An Uneasy Triangle. *Lexikos* 4 (AFRILEX-series 4:1994): 204-214.

Lyons, J. 1968. *Introduction to Theoretical Linguistics.* Cambridge: Cambridge University Press.

Mini, B.M. et al. *The Greater Dictionary of Xhosa.* Volume 2. (Unpublished).

Mtintsilana, P.N. 1990. *Lexicographic Work in Xhosa.* Pretoria: Human Sciences Research Council.

Mtuze, P.T. 1992. A Critical Survey of Xhosa Lexicography 1772-1989. *Lexikos* 2 (AFRILEX-series 2:1992): 165-177.

Pahl, H.W. et al. 1989. *The Greater Dictionary of Xhosa.* Volume 3. Alice: University of Fort Hare Press.

Reader's Digest. 1987. *Universal Dictionary.* Cape Town: The Reader's Digest Association Limited.

Rubin, J. 1984. Bilingual Education and Language Planning. Kennedy, C. (Ed). 1984. *Language Planning and Language Education.* London: George Allen and Unwin.

Smith, N. and D. Wilson. 1979. *Modern Linguistics, The Result of Chomsky's Revolution.* Hammondsworth: Penguin Books.

Stockwell, R.P. 1977. *Foundations of Syntactic Theory.* Englewood Cliffs: Prentice-Hall.

Welmers, Wm. E. 1973. *African Language Structures.* Berkerly and Los Angeles: University of California Press.