
Lemmatisation of Adjectives in Sepedi

R.H. Gouws, *Department of Afrikaans and Dutch,
University of Stellenbosch, Stellenbosch, South Africa* and
D.J. Prinsloo, *Department of African Languages,
University of Pretoria, Pretoria, South Africa*

Abstract: One of the great challenges to compiling better dictionaries for the African languages is to develop sound strategies and procedures for planning the structure of the dictionaries. In this regard all the structural components of a dictionary, including the macrostructure, microstructure, mediostructure and access structure, come into play. Most dictionaries for African languages, including Sepedi dictionaries, fail even at this level. In this article the planning of especially the macrostructure in respect of one lexical category which has been unsatisfactorily treated in Sepedi dictionaries, namely the adjective, will be attempted. Secondly the lemmatisation of adjectives in six Sepedi dictionaries will be critically evaluated. This will be done with the emphasis on various metalexicographical aspects.

Keywords: LEMMATISATION, ADJECTIVES, SEPEDI, MACROSTRUCTURE, USER-PERSPECTIVE

Opsomming: Lemmatisering van adjektiewe in Sepedi. Een van die grootste uitdagings vir die samestelling van beter woordeboeke vir die Afrikatale is om gepaste strategieë en prosedures vir struktuurbeplanning te ontwikkel. Alle struktuurkomponente van 'n woordeboek, insluitende die makro- en mikrostruktuur asook medio- en toegangstrukture is hier ter sprake. Die meeste woordeboeke vir Afrikatale, Sepedi ingesluit, faal in hierdie opsig. In hierdie artikel sal gepoog word om 'n uiteensetting te gee van die wyse waarop die makrostruktuur ten opsigte van die adjektief beplan moet word. Die adjektief is maar een van die kategorieë op hierdie vlak wat onbevredigend in bestaande woordeboeke hanteer is. Tweedens sal die lemmatisering van adjektiewe in ses Sepedi woordeboeke krities beskou word teen die agtergrond van verskeie metaleksikografiese aspekte.

Slutelwoorde: LEMMATISERING, ADJEKTIEWE, SEPEDI, MAKROSTRUKTUUR, GEBRUIKERSPERSPEKTIEF

Introduction

According to Wiegand (1989: 251) lexicography is a practice aimed at the production of dictionaries in order to activate another practice, i.e. the cultural

practice of dictionary use. Any lexicographer compiling a dictionary has the obligation to present the contents of the dictionary in such a way that it will lead to the cultural practice of dictionary use. This can only be achieved if the construction of the specific dictionary adheres to the user-perspective by taking not only the linguistic needs but especially also the reference skills of the intended target user into account. User-friendliness in dictionaries implies that the contents of the dictionary is made as accessible to the user as possible. Attempts to enhance the retrievability of information are often impeded by a high degree of textual condensation. The utilisation of structural markers and other methods to assist the target user in his endeavour of reaching the desired data-presentation means that the internal search route has to be indicated quite clearly. Although this is an important facet of dictionaries, an improvement of the internal search route is not the only way to ensure a better retrievability of information. The macrostructure remains the main access structure of any dictionary with a strictly alphabetical ordering system. Lexicographers too often neglect the importance of a well-designed macrostructure as a functional component of the total linguistic contents of a dictionary by restricting their attempts to enhance user-friendliness to the microstructural level.

The first step towards the improvement of the lexicographic standard of dictionaries for African languages must be to do the groundwork right. Dictionaries are instruments of linguistic and communicative empowerment and therefore lexicographers have to make sure that their intended target users receive an optimal linguistic presentation. To achieve this goal every lexicographer has to rely on a sound theoretical knowledge, and the compilation of every dictionary has to be preceded by the formulation of a business plan, adhering to the aims of the typological criteria of that specific dictionary, and aimed at the specific needs and reference skills of a well-defined target user. This business plan has to be rooted in a general theory of lexicography. According to Wiegand (1984: 14-15) one of the components of a general theory of lexicography is the theory of organisation. This includes all the activities leading to the drawing up of a dictionary plan — that all-important activity that has to precede the compilation of each and every dictionary. The position of the target user may never be underestimated when compiling a dictionary or when drawing up the dictionary plan. Dictionaries are compiled to be used and therefore the target user should be placed in a position where he/she can utilise a dictionary for the successful retrieval of linguistic data.

The traditional, and often haphazard, approach according to which words were entered into a dictionary "as they cross the compiler's way" can no longer be justified. The user-perspective which determines the selection, presentation and treatment of lexical items compels the lexicographer to include those lexical items in the macrostructure that can contribute to the aims of the typological category to which the specific dictionary belongs. The way in which macrostructural elements are presented should also reflect their linguistic status. When dealing with a specific lexical category, the way in which these items

have to be lemmatised, has to be determined on linguistic grounds. The lexicographer has to do an exhaustive analysis of the phenomenon by firstly breaking it down into all its combinations and permutations. Once the compiler is satisfied that he has covered the full scope as viewed from the living language and not only the grammar book, he may start planning how to lexicographically treat the issue within crucial parameters such as the target user's needs, affordability of the dictionary, proper presentation and treatment of the lemma, decisions regarding the data categories to be given, etc. Apart from acquainting himself with sound basic lexicographic principles and practice, he has to study the problematic aspects that the African languages have in common as well as problematic aspects unique to a specific language.

It will be argued that in respect of the adjectives, most dictionaries fail to answer the questions most likely to be asked by their target users, who are usually defined as scholars and students who wish to learn the language. This is due to the lack of a proper needs assessment as part of the overall theory of organisation.

The presentation of adjectives

In the six Sepedi dictionaries used in this present survey, the extremes, with regard to the lemmatisation of adjectives, lie between the *Klein Noord-Sotho woordeboek* where only two forms of a specific adjective are entered into the dictionary without proper guidelines in the front matter on the one hand, and *Sediba* where all possibilities namely nine for each stem are included as lemmas in the central word list on the other.

In planning the macrostructure for a specific lexical category the first step will be to determine whether a limited or an unlimited number of lexical items, i.e. words or stems, are dealt with. The terms "limited" and "unlimited" will be used in a rather oversimplified way. Say, for example, that nouns, verbs, reflexive forms of verbs, etc. are unlimited in that an infinite number of such forms occur while subject concords are limited in that there is only a maximum of 15.

So, in respect of the adjective, the first step will be to determine whether the number of adjectives are limited or unlimited. Only about 30 adjectives of reasonable frequency, listed under (1), occur in Sepedi.

(1)

-bedi	"two"
-raro	"three"
-ne	"four"
-hlano	"five"

-so	"black"
-šweu	"white"
-tala	"green/blue"
-hubedu	"red"
-hwibidu	"red"
-sotho	"brown"
-sehla	"yellow /tawny /grey"
-koto	"thick"
-sese	"thin/narrow"
-thata	"hard/difficult"
-golo	"big"
-nyane	"small"
-telele	"long/tall"
-kopana	"short"
-be	"bad/evil"
-fsa	"new/young"
-swa	"new/young"
-ntši	"many"
-ngwe	"other/some"
-tona	"male"
-tala	"old"
-kaaka	"this big"
-botse	"beautiful/good"
-bose	"delicious/tasty/nice"
-bjalo	"such"
-kae?	"how many?"
-bjang?	"what kind/sort of?"

(Note in passing that some unusual words qualify as adjectives in Sepedi, for example the numbers 2, 3, 4, and 5 as well as the question words *kae?* and *bjang?*)

One of the issues on which the lexicographer has to make a decision is whether it will satisfy the needs of the target user if these adjectives under (1) were lemmatised in that form. What must be kept in mind when answering this question is the lack of typological diversity in Sepedi lexicography. Consequently, the target users of these dictionaries are defined as students and scholars with the inclusion of inexperienced learners. For these users the outer access structure has to provide a direct route to the item they are searching. The typical item they will search, will be words encountered in written texts or oral conversations. None of the adjectives as they are listed under (1), will be found in Sepedi literature. This is due to the fact that these adjectives always have to

take the nominal prefixes of the different noun classes. Compare (2).

(2) **-golo "big/important"**

Class:

1.	monna yo mogolo	monna "man"
2.	batho ba bagolo	batho "people"
3.	mohlare wo mogolo	mohlare "tree"
4.	mebotoro ye megolo	mebotoro "cars"
5.	lesogana le legolo	lesogana "young man"
6.	mahlo a magolo	mahlo "eyes"

The typical target user of the Sepedi dictionaries under discussion, who encounters any occurrence of the adjective *-golo*, will find this lexical item used as the stem of a complex form in which the item *-golo* is preceded by a prefix. This confronts the lexicographer with a dilemma. Pursuing a lexical-based approach to the compilation of the macrostructure (cf. Gouws 1991), the lexicographer will have to include lexical items like stems and affixes in the macrostructure if they have a productive occurrence in real language use. Hausmann and Wiegand (1989: 337) also argue in favour of the fact that all lexical units, including e.g. affixes and other elements of word-formation, may be lemmata. This would mean that lexical items like the stem *-golo* and the prefixes *mo-*, *ba-*, etc. should be included as lemmas in a Sepedi dictionary. The dilemma of the lexicographer is that the reference skills of the target user of the dictionaries under discussion may not equip the user with the expertise to apply the necessary word-formation rules in order to retrieve information about an adjective like *mogolo* from merely consulting the articles of the sublexical lemmas (cf. Gouws 1989) *mo-* and *-golo*. Theoretical soundness and practical realities oppose each other and the lexicographer has to make a difficult decision regarding the forms to be lemmatised.

When deciding on which form to include as macrostructural component, a lexicographer has to consider the theoretical status attributed to that form. According to Hausmann and Wiegand (1989: 329) lemmatisation refers to "the selection of one single morphological form whose function in the macrostructure is to represent the total set of grammatical and morphological forms of the linguistic sign treated in the microstructure". This implies that one lemma sign does not necessarily represent only one lexeme or only one morphological form. Dictionaries usually opt on a systematic basis for one type of item to be lemmatised, e.g. the first person singular form of a verb. Although the treatment is aimed at that lemma sign, it applies to other forms of the lemma as a member of the ordered set of items constituting the treatment units of the dic-

tionary as well. Hausmann and Wiegand (1989: 329) also point out that the inclusion of all irregular forms in the macrostructure is rare.

Adhering to the above-mentioned notion of the lemmatisation of one selected morphological form representing a whole set of forms, a lexicographer can be led to the point where the lemmatisation of adjectives in Sepedi dictionaries does not confront him with any problems. This will imply that only the stem form will be lemmatised and the dictionary user will have to rely on his own linguistic intuition to find the desired information and to apply it to complex words. As noted above, such a lemmatisation system will impede access to the presented data because the lemma sign will not represent a form that can be related to the words found in Sepedi literature. This will characterise the dictionaries as extremely user-unfriendly.

Contrary to the belief that only the stem should be lemmatised, it could also be argued that the complex adjectives consisting of a stem and a prefix are not irregular forms but rather the regular forms of the adjective with the stem as an item which is nonexistent as independent form. Such a word-based approach will not make provision for the lemmatisation of sublexical items like stems, but only for the inclusion of words as lemmas. This will lead to the lemmatisation of all the occurrences of the complex adjectives consisting of e.g. *-golo* plus a prefix. Such a complete list for the different classes will look like column 2 under (3):

(3)

Class	Column 1	Column 2
1.	mogolo	mogolo
2.	bagolo	bagolo
3.	mogolo	---
4.	megolo	megolo
5.	legolo	legolo
6.	magolo	magolo
7.	segolo	segolo
8.	kgolo	kgolo
9.	kgolo	---
10.	kgolo	---
14.	bogolo	bogolo
15.	gogolo	gogolo
16.	gogolo	---
17.	gogolo	---
18.	gogolo	---
	-----	-----
TOTAL:	15	9

If provision for each noun class is to be made, the cost in terms of macrostructural redundancy will be fairly severe. In principle, 15 times 30 = 450 articles, only to make provision for the adjectives in Sepedi.

This number can immediately be reduced to nine since classes 1 and 3, 8-10 and 15-18 respectively take similar forms. However, 9 times 30 still renders a large number of 270 possibilities. The crucial issue will be to maintain a delicate balance between user-friendliness and the possibility of redundancy getting out of hand which in turn directly effects economy and affordability of the dictionary. In simple terms it means that if all of the 270 possibilities are to be accommodated, the dictionary will be very user-friendly since no knowledge of the grammar will be presupposed and all adjectives could be found under the first letter, e.g. *mogolo* under *m-*, *segolo* under *s-*, etc. However, it could be very redundant. This problem could once again activate a tension between the dictionary and the dictionary-using public.

Economy efforts compel the lexicographer to employ space-saving mechanisms — like the lemmatisation of fewer forms. This leads to the professionalisation of lexicography and a high degree of textual condensation. It becomes increasingly difficult for the lay dictionary user to understand this professionalised instrument and to use it successfully. Hausmann (1989: 13) discusses this problem and refers to this conflict between dictionary and user as a conflict between *dictionary culture* and *user-friendliness*. Hausmann sees user-friendliness as the adaptation of lexicography to society whereas dictionary culture is the adaptation of society to lexicography. This means that user-friendliness demands that the contents and presentation of a dictionary should be determined by the needs and expertise — or lack thereof — of society. Dictionary culture means that society has to be educated to utilise more sophisticated dictionaries.

Possible solutions in Sepedi dictionaries

One extreme solution to the problem could be to reduce column 2 under (4) from nine possibilities to only TWO as in column 4.

The other extreme would be to enter the full range of 270 possibilities into the dictionary with exhaustive treatment in each case, which will of course be very user-friendly but extremely redundant. Lexicographers have to endeavour to make these extremes more viable. The major challenge will be to make the first extreme, namely to lemmatise only two forms as under (4) column 4, more user-friendly.

(4)

Class	Column 1	Column 2	Column 3	Column 4
1.	mogolo	mogolo	-golo	-golo
2.	bagolo	bagolo	-golo	---
3.	mogolo	---	-golo	---
4.	megolo	megolo	-golo	---
5.	legolo	legolo	-golo	---
6.	magolo	magolo	-golo	---
7.	segolo	segolo	-golo	---
8.	kgolo	kgolo	kgolo	kgolo
9.	kgolo	---	kgolo	---
10.	kgolo	---	kgolo	---
14.	bogolo	bogolo	-golo	---
15.	gogolo	gogolo	-golo	---
16.	gogolo	---	-golo	---
17.	gogolo	---	-golo	---
18.	gogolo	---	-golo	---
	-----	-----		-----
TOTAL:	15	9		2

Utilising the front matter

A possible way of coping with this problem is to utilise the front matter of the dictionary by including easy to read guidelines, e.g.:

“In this dictionary adjectives are entered on the stem, e.g. *mogolo* in an example such as *monna yo mogolo* ‘a big/tall/important man’ must be looked up under word minus stem, that is *mogolo* – *mo* = *-golo*.” Thus the complete table of guidelines would be as in (5):

(5)

1 and 3	mogolo	look up under	-golo
2.	bagolo	look up under	-golo
4.	megolo	look up under	-golo
5.	legolo	look up under	-golo
6.	magolo	look up under	-golo
7.	segolo	look up under	-golo
14.	bogolo	look up under	-golo
15-18	gogolo	look up under	-golo

The form for classes 8, 9 and 10, *kgolo*, will be lemmatised as *kgolo* and is no problem.

Within a target user community with a well-developed dictionary culture this approach could surely be defended. Dictionaries have to be regarded as carriers of texts (cf. Wiegand 1996). In a dictionary as a text carrier that displays a typical textual book-structure, the central word list is a compulsory text. All functional text parts preceding this central word list constitute the front matter and all the functional text parts following the central word list constitute the back matter of the dictionary (cf. Hausmann and Wiegand 1989: 330-331). Besides the central word list there is only one other obligatory text, i.e. the text in the front matter containing the user's guidelines. Because this is an obligatory text, the lexicographer may include information in this text which will assist the user to achieve an optimal retrieval of information from the central word list. When adjectives are treated in Sepedi dictionaries, there should, from a metalexigraphic perspective, in principle be no objections to a limited lemmatisation of this word class if the front matter contains a text with user's guidelines in which a sound and systematic explanation of this word class is given.

Once again, however, the potential conflict between user-friendliness and dictionary culture has to be taken into account. Hartmann (1989: 103) argues that an analysis of user's needs should precede dictionary design. The lexicographer of a Sepedi dictionary should allow the outcome of a needs and reference skills analysis to determine a variety of characteristics of the dictionary. One aspect to be considered by the lexicographer is whether the typical target user is in the habit of utilising the texts in the front matter to improve his dictionary using skills or his access to the presented information.

Unfortunately lexicographers may seldom rely on the willingness or habit of their target users to utilise a text that does not form part of the central word list. Therefore Busane (1990) is in the right when he says that dictionary users are not known for consulting the guidelines to the dictionary, they want to find what they need instantly without referring to grammatical rules and guidelines in the front matter or even guidelines within the dictionary itself.

Alternative possibilities

The main weakness of the other extreme, namely to lemmatise all the adjectives, could be combated by attempts to reduce redundancy by, among others, (a) reduction based on frequency-of-use, (b) shorter articles including less data categories and (c) cross-references.

(a) Reduction based on frequency-of-use

The compiler could decide to omit the adjective due to the fact that the overall

count of eleven occurrences for classes 15-18 under (6) is very low in comparison to the rest. This is especially the case for other adjectives which are less frequently used than *-golo*.

(6)

1	2	3	4	5	6	7	8	9
Class	'big'	Freq.	Sediba	Popular	Klein	Pukuntšu	Shuters	New
1 and 3	mogolo	611	mogolo	---	---	mogolo	mogolo	mogolo
2	bagolo	328	bagolo	---	---	bagolo	bagolo	---
4	megolo	70	megolo	megolo	---	---	megolo	---
5	legolo	243	legolo	---	---	legolo	legolo	---
6	magolo	136	magolo	magolo	---	magolo	magolo	magolo
7	segolo	161	segolo	segolo	---	segolo	segolo	segolo
8-10	kgolo	597	kgolo	kgolo	kgolo	kgolo	kgolo	kgolo
14	bogolo	241	bogolo	bogolo	---	bogolo	bogolo	bogolo
15-18	gogolo	11	gogolo	---	---	gogolo	---	gogolo
	-golo	0	---	golo	-golo	-golo	---	golo

The ways in which adjectives have actually been lemmatised in six Sepedi dictionaries will be evaluated with reference to (6).

Column 1 gives the noun class or classes related to the specific form of the adjective, column 2 the adjectives for classes 1-18.

In column 3 the overall frequency count on a one million corpus, compiled from fifty different books and magazines, is shown, followed in columns 4-9 by an indication of the inclusion or omission of the adjectives in the Sepedi dictionaries in question. It is clear from column 3 that this adjective is in principle highly used in Sepedi. (A total count for all the classes is 2398 which means that it is used more than 40 times on average in every single Sepedi book or magazine.) Furthermore it is clear that the forms *mogolo* (classes 1 and 3), *kgolo* (classes 8-10) and *bagolo* (class 2) are the most highly used.

As indicated in column 4, all the relevant forms are entered in *Sediba*, which represents one of the extremes. This is more or less as good as it can be in respect of user-friendliness.

According to column 5, all the relevant forms, except the forms for classes 1 and 3, 2, 5 and 15-18, are given in the *Popular Northern Sotho Dictionary*. In addition, the stem *golo* is given, but as a word, that is, without the hyphen indicating its status as a sublexical lemma. This is unacceptable, especially in view of its high frequency of use in classes such as 1-3.

As shown in column 6, the compilers of the *Klein Noord-Sotho woordeboek* opted for the other extreme, namely to enter only the form *kgolo* for classes 8-10 and the stem *-golo* for the rest. (Compare column 4 under (4) once again.) In fact, only *-golo* was entered and treated, while *kgolo* was entered with a cross-

reference to *-golo*.

According to column 7, *Pukuntšu* gives all the relevant forms, with the exception, for no apparent reason, of class 4. Also entered is *-golo*, properly marked as a stem.

As indicated in column 8, all the relevant forms are entered in the *Shuters New Sepedi dictionary*, with the exception of class 15-18 which was omitted on the basis of low frequency. As shown in the case of *Sediba* in column 4, it is unnecessary to enter the stem form *-golo* as well, since all the derivations have been covered.

Finally, in column 9 the entries for the *New English-Northern Sotho Dictionary* are given with the forms for classes 2, 4 and 5 missing and *golo* entered as a word instead of a stem.

(b) Shorter articles including less information categories

In addition to attempting reduction based on frequency-of-use, shorter articles could be employed. Articles could be shortened in various ways. A decrease of the data types would also decrease the density of information. If this is done on the basis of a needs analysis which results in the omission of redundant or less functional data categories, this option could lead to an increase in the users' comprehension. However, the articles can also be shortened by a process of textual condensation that does not omit data categories but retains them although in a more condensed presentation. Textual condensation, accompanied by a high degree of information density, results in a more complex microstructural presentation which impedes the retrieval of information and the successful interpretation of the articles. According to Kühn (1989: 112) the use of a dictionary has to be understood as a communicative act. The lexicographer has to endeavour to improve the quality of this communicative act. In a dictionary aimed at scholars, students and learners, textual condensation will definitely be detrimental for the user when employing the dictionary in a communicative act. The inclusion of all adjectives as lemma signs is a user-friendly option. However, if this is accompanied by a treatment that omits certain data categories or that condenses the presented data, the question arises whether it would not have been better to utilise the available space for a more extensive treatment of fewer lemmas.

(c) Cross-references

A lexicographic procedure that has not yet had an optimal employment in South African dictionaries, is the dictionary-internal mediostructure. According to Wiegand (: 196: 11) the dictionary-internal mediostructure interconnects the knowledge elements represented in different sectors of the dictionary on several levels of lexicographic description. Wiegand (1996: 11) continues:

A lexicographer *refers* the potential user from a *reference position* giving the *reference item* or other *reference transmitting items* to the *reference address*, which possibly provides access to the lexicographic data relevant for obtaining the user's objective. Thus, a *reference relation* is established either between the reference item or other reference transmitting items to one or more reference address(es).

One of the biggest advantages of the effective utilisation of a dictionary-internal mediostructure is that precious space can be saved by, for example, giving an exhaustive treatment of one entry with cross-references from the other skeleton entries. This could be regarded as user-unfriendly in a different way, as is the case in (5) where the user has to consult and rely on guidelines given in the nonalphabetical section. However, if the reference address is a lemma in the central word list, the system of cross-referencing can enhance the text-internal cohesion. This can also lead the user to experience the lexicon as network of relations.

In Sepedi dictionaries the employment of a procedure of dictionary-internal mediostructural relations will compel the lexicographer to give an explicit explanation of the system in the front matter of the dictionary. The application of a system of cross-referencing should be done in such a simple and explicit way that even the user who does not consult the front matter has to be able to follow the reference route and to retrieve the necessary information from the treatment of the reference address. This would mean that all the various occurrences of an adjective can be lemmatised but these lemmas will receive a limited lexicographic treatment and will primarily be used as reference items filling the reference position. Besides grammatical information, e.g. an indication of the nominal prefix and the specific noun class, the treatment will consist of an indication of the reference address. This reference address could be the stem which is the salient component of each adjective.

Conclusion

The lemmatisation of adjectives may no longer be done in an arbitrary way. A detailed analysis of the problems and possible solutions is a prerequisite for the compilation of a proper macrostructure. Each and every aspect should be subjected to a similar analysis before one could think of tackling the microstructure. In this regard the lexicographer has to rely on the results of metalexicographical research.

References

1. General

- Busane, M. 1990. Lexicography in Central Africa: The User Perspective with Special Reference to Egypt. Hartmann, R.R.K. (Ed.). 1990: 19-35.

- Gouws, R.H. 1989. *Leksikografie*. Cape Town: Academica.
- Gouws, R.H. 1991. Toward a Lexicon-based Lexicography. *Dictionaries* 13: 75-90.
- Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings*. Tübingen: Max Niemeyer.
- Hartmann, R.R.K. 1989. Sociology of the Dictionary User: Hypotheses and Empirical Studies. Hausmann, F.J. et al. (Eds.). 1989: 102-111.
- Hartmann, R.R.K. (Ed.). 1990. *Lexicography in Africa*. Exeter Linguistic Studies 15. Exeter: University of Exeter Press.
- Hausmann, F.J. and H.E. Wiegand. 1989. Component Parts and Structures of General Monolingual Dictionaries: A Survey. Hausmann, F.J. et al. (Eds.). 1989: 328-360.
- Hausmann, F.J. 1989. Die gesellschaftlichen Aufgaben der Lexikographie in Geschichte und Gegenwart. Hausmann, F.J. et al. (Eds.). 1989: 1-19.
- Hausmann, F.J. et al. (Eds.). 1989. *Wörterbücher. Ein internationales Handbuch zur Lexikographie / Dictionaries. An International Encyclopedia of Lexicography. / Dictionnaires. Encyclopédie internationale de lexicographie*. Berlin: Walter de Gruyter.
- Kühn, P. 1989. Typologie der Wörterbücher nach Benutzungsmöglichkeiten. Hausmann, F.J. et al. (Eds.). 1989: 111-127.
- Wiegand, H.E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Ed.). 1984: 13-30.
- Wiegand, H.E. 1989. Der gegenwärtige Status der Lexikographie und ihr Verhältnis zu anderen Disziplinen. Hausmann, F.J. et al. (Eds.). 1989: 246-280.
- Wiegand, H.E. 1996. Über die Medialstrukturen bei gedruckten Wörterbüchern. Zettersten, A. and H. Pedersen (Eds.). 1996: 11-43.
- Zettersten, A. and H. Pedersen (Eds.). 1996. *Symposium on Lexicography VII*. Tübingen: Max Niemeyer.

2. Dictionaries

- Kriel, T.J. 1985. *New English-Northern Sotho Dictionary*. Johannesburg: Educum.
- Kriel, T.J. 1988. *Popular Northern Sotho Dictionary*. Pretoria: J.L. van Schaik.
- Kriel, T.J. and E.B. van Wyk. 1989 *Pukuntšu*. Pretoria: J.L. van Schaik.
- Lombard, D.P. et al. 1991. *Sediba*. Pretoria: Via Afrika.
- Prinsloo, D.J. and B.P. Sathekge. 1996. *Shuters New Sepedi Dictionary*. Pietermaritzburg: Shuter and Shooter.
- Ziervogel, D. and P.C. Mokgokong. 1979. *Klein Noord-Sotho woordeboek*. Pretoria: J.L. van Schaik.