The Use of an Improved Access Structure in Dictionaries*

Rufus H. Gouws, Department of Afrikaans and Dutch, University of Stellenbosch, Stellenbosch, Republic of South Africa

(rhg@akad.sun.ac.za)

Abstract: Dictionary consultation procedures often fail because the structure of a dictionary impedes the user from sufficient access to the data. Modern-day metalexicography has identified the access structure as one of the most important structural components of a dictionary. The access structure represents the search route a user follows to reach the required data in order to retrieve the necessary information. This paper focuses on a few aspects of the access structure. It is argued that a frame structure and an improved use of outer texts can enhance access to the lexicographic data. Different macrostructural ordering procedures are discussed and it is suggested that the lexicographers of bilingual dictionaries should avoid multiple niching. The presentation of inner texts in clearly identifiable text blocks can assist the user in rapidly reaching the desired search zone. Although the emphasis is on bilingual dictionaries many of the suggestions also apply to monolingual dictionaries.

Keywords: METALEXICOGRAPHY, LEXICOGRAPHY, DICTIONARY, BILINGUAL, ACCESS STRUCTURE, FRAME STRUCTURE, MULTIPLE NICHING, NESTING, INNER TEXTS, OUTER TEXTS

Opsomming: Die gebruik van ‘n verbeterde toegangstruktuur in woordeboeke. Naslaanpogings in woordeboeke misluk dikwels omdat die woordeboekstruktue die gebruiker van doelvullende toegang tot die data weerhou. In die hedendaagse metaleksikografie word die toegangstruktuur as een van die belangrikste strukturele komponente van ‘n woordeboek beskou. Die toegangstruktuur verteenwoordig die soekroete wat ‘n gebruiker volg om die verlangde data te bereik en die nodige inligting te ontsluit. Hierdie artikel fokus op enkele aspekte van die toegangstruktuur. Daar word aangetoon dat ‘n raamstruktuur en ‘n verbeterde gebruik van buitetekste die toegang tot leksikografiese data kan verbeter. Verskeie makrostrukturele ordeningsmetodes word bespreek en daar word voorgestel dat leksikograwe van vertalende woordeboeke moet probeer om veelvuldige nisting te vermy. Die plasing van binnetekste in duidelik identifieerbare teksblokke, kan die gebruiker help om ‘n verlangde soeksonte vinnig te bereik. Alhoewel die klem op tweetalige woordeboeke is, geld baie van die opmerkings ook eentalige woordeboeke.

Sleutelwoorde: METALEKSIKOGRAFIE, LEKSIKOGRAFIE, WOORDEBOEK, TWEE

* This article is a slightly adapted version of a paper presented at the Second International Conference of the Asian Association for Lexicography, held at Yonsei University, Seoul, 8-10 August 2001.

1. Introduction

The success of a bilingual dictionary depends largely on the success, which the typical target user achieves during a normal dictionary consultation procedure. Too often these procedures fail because, although bilingual dictionaries contain all the relevant data, the structure of the dictionary and the presentation of the data is of such a nature that the target user cannot find the data or retrieve the needed information. The rapid and unimpeded access of the user to the relevant data presented in the dictionary has to be regarded as a prerequisite for a successful lexicographic product in a user-driven approach.

Recent research in the field of metalexicography has focused on the structure of dictionaries. One of the structural components identified in this research (cf. Hausmann and Wiegand 1989), is the access structure. The access structure can be regarded as the search route a dictionary user follows during a dictionary consultation procedure. A distinction is made between the outer access structure, which leads a user up to the lemma sign introducing the article from which the needed information is to be retrieved, and the inner access structure which guides the user within a dictionary article to the search zone in which the relevant data is presented. In this paper the importance of both the outer and the inner access structures will be emphasised. However, the need will also be stressed for a more sophisticated access structure and for the use of outer texts to enhance the rapid and successful access to the data presented in dictionaries.

2. The frame structure

A user-driven approach in lexicography compels the lexicographer to use all possible means at his/her disposal to improve the quality of dictionary use and to ensure an optimal retrieval of information. Successful information retrieval can only be achieved if the data is accessible. Accessibility of data is not only determined by the use of a well-defined access structure leading the user to a venue within the word list structure of the dictionary but it is also determined by a functional positioning of the data within the word book structure of the dictionary and by employing different texts in the dictionary to explain and explicate the contents of that dictionary (cf. Hausmann and Wiegand 1989: 330, 333).

Utilising a word book approach a dictionary can display a so-called frame structure (cf. Kammerer and Wiegand 1998). A dictionary is regarded as a carrier of different text types and the different texts and text types are accommodated in various positions within the dictionary. According to the frame structure approach a dictionary contains three textual focal points, i.e. the front
matter, the central list and the back matter. The central list contains the article stretches, i.e. the presentation of the articles grouped according to the alphabetical ordering system in terms of the different alphabet letters A-Z, in which all the articles are accommodated. This is the section of a dictionary which is the most typical venue for general dictionary consultation procedures. Consequently this textual component has to display a structure which allows the target user quick and unimpeded access to the required data. The outer access structure will guide a user to a specific article stretch and within that article stretch to a specific lemma sign. From there the inner access structure leads the user via an article internal route to the relevant search zone and the specific data type. The front matter section contains all the texts preceding the central list and the back matter section contains all the texts following the central list. The texts in the front and back matter of a specific dictionary constitute the outer texts of that dictionary. Neither the front nor the back matter are functional parts of a dictionary, although both may contain individual texts which have specific functions in the dictionary (cf. Gouws 1999). A frame structure qualifies a dictionary as a text type carrier or a "big text" because the front matter, central list and back matter contain a variety of text types.

The dictionary conceptualisation plan of any new lexicographic endeavour has to make provision for a data distribution programme according to which the data which the lexicographer wishes to include can be distributed to all the focal points of the planned dictionary. Its application results in the introduction of a data distribution structure. Traditional dictionaries have often utilised the front and back matter to present outer texts. However, in many dictionaries these were not functional texts and contributed very little to the genuine purpose of the dictionary.

Due to the functional use of both the central list and the outer texts a dictionary can assist with a much more comprehensive transfer of data than before, e.g. the inclusion of various lists of words in the back matter. The inclusion and distribution of more data in different structural focal areas of a dictionary necessitate clear and unambiguous access to the data. The access structure of a dictionary has to be enhanced to ensure the success of a dictionary consultation procedure performed by a typical target user of the specific dictionary. Access structures are not only needed to guide the user to the article stretches in the central list but also to assist the user in utilising and retrieving information from the texts included in the front and back matter components.

### 2.1 Outer texts

According to Hausmann and Wiegand (1989: 331) each dictionary should contain two compulsory texts, i.e. the central list which presents the lemmata as macrostructural entries and primary treatment units and an outer text, usually presented in the front matter, with the users' guidelines. In this outer text the editorial system of the dictionary has to be explained in order to familiarise the
target user with the structure and contents of the dictionary and with the presentation of data. This text already has an access function because it not only informs the user of the data collection in the dictionary but also explicates the positioning and arrangement of data types. This allows the user a more rapid dictionary consultation procedure. A typical component of such a users' guidelines text is a sample article taken from the central list to illustrate the default article structure, microarchitecture and search zones. With such an outer text the access possibilities surpass those of the traditional outer and inner access structure.

The notion of a frame structure, as developed by Kammerer and Wiegand (1998), creates the opportunity for a more comprehensive application of front and back matter texts (cf. Gouws 2001), which further improves the accessibility of dictionaries. Frame structures of different levels can be identified. According to Kammerer and Wiegand (1998) the front and back matter texts are not immediate constituents of the front or back matter respectively but rather of the dictionary as a carrier of texts. A front matter text functioning as an immediate constituent of the dictionary as a whole, e.g. a users' guidelines text or a front matter text presenting the table of contents of the dictionary, can be regarded as a primary outer text. Bergenholtz, Tarp and Wiegand (1999: 1777-1778) argue convincingly in favour of a distinction between integrated and nonintegrated outer texts. An integrated outer text is one integrated into and helping to achieve the genuine purpose of the specific dictionary. It quite often has a direct link with the central list, e.g. a dictionary displaying both single and synopsis articles can assist the user by listing the lemmata heading the synopsis articles in a back matter text. Such a text explicates the heterogeneous article structure and creates another access possibility for the user. In this way integrated outer texts, both in the front and the back matter, can be employed to create polyaccessible dictionaries. In a bilingual dictionary the treatment of lemmata representing culturally bound lexical items could best be done in synopsis articles which allows a more comprehensive treatment than that presented in default articles. A back matter text listing all the culturally bound items treated in the central list in synopsis articles enhances the access to these items in the central list.

The relevance and functional value of an outer text does not depend on its integration into the genuine purpose of the dictionary. Gouws (forthcoming) indicates that an outer text can have its own genuine purpose which does not necessarily correspond with the genuine purpose of the dictionary or its integrated outer texts. These unintegrated outer texts can play an important role in the overall transfer of data in a lexicographic text and their selection and presentation may not be done in an arbitrary way. The data distribution structure of a dictionary determines the type of data to be presented in outer texts and the target user of a dictionary needs to be made aware of their contents and gain access to the data in these texts. The average dictionary user is often not in the habit of utilising the outer texts of a dictionary. The functional value and
relevance of outer texts should not only be brought to the attention of the target users but lexicographers should endeavour to promote optimal access to these texts. This demands an innovative approach towards the access structure. The access structure should not be biased by being directed only at the central list. As an integral component of the word book structure, it should establish a search route to all the venues identified by the data distribution programme for the accommodation of lexicographic data. In this regard the table of contents plays an important role in giving a systematic account of the different texts included in the dictionary.

A text presenting the table of contents has a wide scope because it is directed at the dictionary as a carrier of text types and gives information regarding the texts functioning within the different focal areas constituting the word book structure. According to Hausmann and Wiegand (1989: 333) a table of contents functions as a special type of access structure, namely a register-external outer access structure, i.e. an outer access structure component which is not presented as part of the central list. Where the typical outer access structure guides a user to a lemma sign, the table of contents includes entries which guide the user to the different components of the frame structure and to the individual texts contained in these components of the dictionary as a big text. Such a table of contents is part of the primary frame structure of the dictionary and can be regarded as a primary outer text. The functional value of a table of contents which increases accessibility may never be underestimated. This text can assist the intended target user in a substantial way.

Accessibility to the outer texts can be achieved in different ways. One access procedure which has not yet been employed enough is the utilisation of secondary outer texts. Many dictionaries contain primary outer texts with their own outer texts, e.g. an explanatory text, a table of contents or a users’ guidelines text directed not at the dictionary as such but only at a specific outer text. These secondary outer texts are not immediate constituents of the dictionary as a big text but rather of the specific outer texts. A primary outer text with its own front and back matter texts constitutes a secondary frame structure (cf. Gouws 2001: 527). A major function of secondary outer texts is to provide access to their primary outer texts and to increase accessibility to data accommodated outside the central list of a dictionary.

3. The external access structure: macrostructural procedures

All dictionaries do not display the same macrostructural ordering procedures. A typical distinction, occurring in dictionaries with an alphabetical ordering, is that between a straight alphabetical ordering and an ordering presenting a sinuous lemma file (cf. Wiegand 1989, Hausmann and Wiegand 1989). These different conventions result in different types of lemmata (cf. Wolski 1989). A dictionary with a straight alphabetical ordering typically displays main lemmata whereas a dictionary with a sinuous lemma file has sublemmata comple-
menting the main lemmata. Sublemmata are ordered in either lemma niches or
lemma nests. In monolingual explanatory dictionaries sublemmata often are
the so-called “unexplained” complex lexical items. Their semantic transparency
motivates their inclusion as sublemmata. This linguistic motivation is comple-
mented by a pragmatic one, i.e. for space-saving purposes. Albeit that they are
not allocated paraphrase of meaning entries, sublemmata can still convey cer-
tain semantic data, especially if they are entered as nested lemmata.

3.1 Niching and nesting

Niched lemmata display a straight alphabetical ordering and the niche main-
tains the alphabetical ordering in terms of the preceding and the following
main lemmata. The deviation in ordering direction, i.e. horizontal as opposed
to vertical, has no influence on the alphabetical ordering principle. This is clear
from the following example, taken from the Afrikaans explanatory dictionary
Verklarende Afrikaanse Woordeboek:

eergie’rig. …
eer’gister. Die dag voor gister; eergisteraand; eergistermiddag; eergistermôre; …
eer’herstel. …

Nesting differs from niching in at least one but often two ways. Lemma nests
deviate from the straight alphabetical ordering. Although a lemma nest some-
times displays an internal alphabetical ordering similar to a niche, the external
ordering, i.e. the alphabetical positioning of the nested lemmata with regard to
the preceding and following vertically ordered lemmata, does not display a
straight alphabetical ordering. This kind of nesting may be referred to as first
level nesting. Cf. the following example taken from the Verklarende Handwoorde-
boek van die Afrikaanse Taal:

dia·mant’ (-e) Harde, kosbare edelsteen, bestaande uit kristallyne koolstof:
(On)geslypte diamante. [G. adamas staal; L. diamas onbedwingbaar; G. damao
mak mak] diamant: ~agtig, ~bedryf, ~boor, ~delwer, ~gruis, ~houdend,
~kenner, ~ontginning, ~ring, ~vorming, ~winning.
dia·mant’ bruilof

A straight alphabetical ordering would have ordered the lemma diamantbruilof
between diamantboor and diamantdelwer. Internally a nest may also present a
deviation from the alphabetical ordering. This is known as second level nesting
and is usually motivated on morphosemantic grounds. Cf. the following exam-
ple from Nasionale Woordeboek:

broei (ge-) wv. 1. op eiers sit en hulle warm hou om hulle te laat uitkom. 2. voortkom,
ontspruit. Daaruit sal onheil –. 3. peins, planne maak. Oor iets –. 4. ontwikkel,
in wording wees. Daar is iets aan die –. 5. hitte ontwikkel, warm word. Die mis,
lug –. 6. warm word deur gisting. Die hooi –. 7. deur spesiale verwarming
vroeër laat bloei of ryp word. 8. warm kry. In die son sit en –. 9. kleintjies
The first two lemmata in the nest are derivatives and the rest are compounds. The first semicolon is a structural marker separating members of the nest on morphological grounds. A partial alphabetical ordering prevails but the nest is characterised by different subgroupings of lemmata. Each subgroup is followed by an entry indicating the relevant subcomment on semantics of the main lemma with which the preceding sublemmata can be associated. This deviation from the alphabetical ordering is determined on semantic grounds. Lemma niching and, especially, lemma nesting involve procedures of textual condensation which can impede rapid and easy access to the presented data.

Where the niching and nesting also lead to the use of place-keeping symbols in a condensed lemma sign it increases the degree of textual condensation (cf. Wolski 1989a: 964).

A straight alphabetical ordering is much more user-friendly and allows more successful access than an ordering displaying niched and nested lemmata. The dictionary typology and the target user should determine the nature of the presentation, ordering and textual condensation. A sinuous lemma file can impede dictionary consultation procedures in a monolingual dictionary and in a bilingual dictionary, bilingual or multilingual, it can create serious access problems, especially where the lemma signs are the targets of textual condensation. Where a lexical item has a frequent occurrence as first stem of a compound, the lemmatisation of these compounds can impede the access procedure if the lexicographer opts for a system of multiple niching instead of first level nesting. First level nesting can be seen in the following example from New Words/Nuwe Woorde:

**con ⋅ sum ⋅ er**

**con ⋅ sum ⋅ er ⋅ ism**
- **n.** (dikw. neerh.) verbruikerisme. **con ⋅ sum ⋅ er ⋅ ist**
- **n.** verbruikerskampvegter; konsumentis, aanhanger/voorstander van die konsumentisme. consumerist a. verbruiker- (samelewing ens.); verbruikers- (etiek, kultuur, ens.).

In the following example from Groot Woordeboek/Major Dictionary the lexicographer opts for multiple niching:

**moon**, (n) maan; maand (poëties); ... ~beam, maanstraal; ~blind, nagblind; ~blindness, nagblindheid; ~buggy, maanbesie, maanbakkie; ~calf, gedrog, misgeboorte; uilskuiken ... ~landing, maanlanding; ~less, donker(maan), sonder maan; a ~less night, donkermaan, 'n donker aand.
Here the niches are constructed by means of lemma parts (cf. Wolski 1989: 365). In the article clusters moon: ~lit, ... ~set and moon: ~shine, ... ~year the niches do not contain a single full lemma. They are introduced by lemma parts in niche external position and the clusters are characterised by condensed lemmata consisting of a place-keeping symbol and the second stem of the respective compound. It is clear that a user looking for translation equivalents of, say, the English words moonrise or moon-watch, should be quite sophisticated in dictionary using skills to ensure rapid access to the lemmata. Multiple niching decreases the accessibility of the external search route and the dictionary consultation procedure. Where bilingual dictionaries opt for a sinuous lemma file this ordering should be restricted to first level nesting or to single niching. Even second level nesting, although it presents a more comprehensive transfer of linguistic data, can create an impediment in the outer access structure. In order to improve the quality of the external access structure the lexicographer should try to refrain from the use of condensed lemmata.

4. The use of inner texts

Inserted texts can be employed to enhance the access to certain macrostructural entries. An inner text, presented in a clearly identifiable frame, can be regarded as a salient text block and search zone which immediately attracts the attention of a user and highlights the search route to the desired lemma sign. Inserted texts can be presented in different ways, e.g. as inner texts within an article or as a text block at the end of an article. The Kenkyusha’s Lighthouse English–Japanese Dictionary makes provision for both these procedures. Inserted inner texts, positioned between two subcomments on semantics of a polysemous lemma, present lemma nests, adhering to an internal alphabetical ordering, of compounds in which the main lemma functions as second component. Because the second stem of a component constitutes the semantic core, these compounds are semantically related to the main lemma. In a bilingual dictionary where the user is not familiar with the source language the clustering of semantically
related items as sublemmata in an inserted inner text could be beneficial to ensure access to the data. Cf. the presentation of the sublemmata hairbrush, nail brush, etc. in the inserted inner text positioned between the first two subcomments on semantics of the article of the lemma sign brush. The article of the lemma sign brush is followed by an inserted text which introduces a cluster of nested articles with sublemmata brush side, brush down, etc., representing particle verbs with the main lemma as first component. The frame in which the article cluster occurs marks a specific search zone and forms an integral part of the access structure. If this system is explained clearly in the users' guidelines text and applied consistently throughout the dictionary it will improve the accessibility of the dictionary.

---

5. The inner access structure: microstructural procedures

Bilingual dictionaries have too often been characterised by a lemmatic bias. This has resulted in an access structure dominated by the external search route and also in an addressing structure where the lemmata function not only as primary but often as the only treatment units. Access to target language entries has often been of an inferior nature. To enhance the access structure a much stronger focus on target language forms such as nonlemmatic addresses and secondary treatment units is needed. This demands a more sophisticated system of typographic and nontypographic structural markers to guide the user on the article internal search route towards secondary treatment units. Lexicographers should consider the use of a well-structured microarchitecture and the employment of unambiguously defined search zones, arranged in clearly distinguishable text blocks to assist the user when negotiating the rapid access structure of a dictionary.
Modern-day lexicography is dominated by the user-perspective and the access structure of any new dictionary has to be user-driven, i.e. it has to be planned in accordance with the user profile of the dictionary and the needs and reference skills of a well-defined target user. Where the speakers of the languages treated in a bilingual dictionary have different cultural backgrounds, the data distribution and presentation will necessarily have to make provision for these differences. The access structure will have to assist users to identify surrogate equivalents rendered in the treatment of culturally bound source language forms. These equivalents will have to be the treatment units of contextual and cotextual items. Surrogate equivalents must be clearly distinguishable from other equivalents by means of structural markers.

6. In conclusion

Access structure models should not be of a language specific nature and must be open enough for the planners of new dictionary projects in various languages and subdomains of bilingual and multilingual lexicography to customise them in terms of the needs of their specific projects. The customisation, however, must ensure an unimpeded access which can result in a dictionary consultation procedure which renders an optimal retrieval of information.

References


Gouws, Rufus H. 1999. Mediostructural Representation, Textual Condensation and User-Orienta-

tion in the WAT X. Lexicographica 15: 4-37.


