
Portals to Knowledge: CD-ROM Encyclopaedias, with Specific Reference to *Microsoft Encarta 99 Encyclopedia*

Phillip Adriaan Louw, *Department of Afrikaans and Dutch,
University of Stellenbosch, Stellenbosch, Republic of South Africa*

Abstract: The close interaction between linguistics and lexicography over the last few decades has had many good effects, not least of which has been the development of a theoretical framework or metalexicography. However, this close link has had an unfortunate result: the marginalisation in the theoretical debate of nonlinguistic dictionaries and especially of the encyclopaedia.

In this article, semantic, lexicographical and pragmatic motivations will be given for a renewed interest in encyclopaedias. The revolution caused by CD-ROM encyclopaedias and the benefits they hold for a large spectrum of users, will be the most important pragmatic motivations addressed. *Microsoft Encarta 99 Encyclopedia* will act as a guide for the exploration of the impact this revolution has had on the ideas of user-friendliness and dictionary accessibility. Finally, a future is envisioned in which the Internet will be the logical culmination of the ultimate potential of computers as accessible, user-friendly reference tools.

Keywords: ACCESS STRUCTURE, ACTIVE CROSS-REFERENCES, CD-ROM ENCYCLOPAEDIA, DICTIONARY ACCESSIBILITY, ENCYCLOPAEDIA, GUIDE STRUCTURES, INNER ACCESS STRUCTURE, INNER SEARCH PATH, INTERNET, LINGUISTIC DICTIONARY, MEDIOSTRUCTURE, MULTIMEDIA, NONLINGUISTIC DICTIONARY, OUTER ACCESS STRUCTURE, OUTER SEARCH PATH, PASSIVE CROSS-REFERENCING, USER-FRIENDLINESS

Opsomming: Poorte na kennis: CD-ROM-ensiklopedieë met spesifieke verwysing na *Microsoft Encarta 99 Encyclopedia*. Die noue interaksie tussen die linguistiek en die leksikografie het in die laaste paar dekades reeds baie goeie gevolge gehad, met die totstandkoming van 'n teoretiese raamwerk of metaleksikografie as seker die prominentste uitvloeisel. Hierdie wisselwerking het egter tot 'n ongelukkige gevolg gelei: die marginalisasie van nielinguistiese woordeboeke en veral ensiklopedieë binne die teoretiese debat.

In hierdie artikel sal semantiese, leksikografiese en pragmatiese motiverings aangebied word vir 'n hernieude fokus op ensiklopedieë. Die revolusie wat deur die CD-ROM-ensiklopedieë veroorsaak is en die voordele wat hulle vir 'n wye spektrum gebruikers inhou, is die belangrikste pragmatiese motiverings wat bespreek sal word. *Microsoft Encarta 99 Encyclopedia* sal optree as gids by die ondersoek na die impak wat hierdie revolusie op die beginsels van gebruikersvriendelikheid en woordeboektoeganklikheid het. Laastens word 'n blik op die toekoms gebied waarin die Inter-

net as die logiese bereiking van die potensiaal van rekenaars as toeganklike, gebruikersvriendelike naslaanbronne in die vooruitsig gestel word.

Sleutelwoorde: AKTIEWE KRUISVERWYSINGS, CD-ROM-ENSIKLOPEDIE, EKSTERNE SOEKROETE, EKSTERNE TOEGANGSTRUKTUUR, ENSIKLOPEDIE, GEBRUIKERSVRIENDELIKHEID, GIDSSTRUKTURE, INTERNE SOEKROETE, INTERNE TOEGANGSTRUKTUUR, INTERNET, LINGUISTIESE WOORDEBOEK, MEDIOSTRUKTUUR, MULTIMEDIA, NIELINGUISTIESE WOORDEBOEK, PASSIEWE KRUISVERWYSINGS, TOEGANGSTRUKTUUR, WOORDEBOEKTOEGANKLIKHEID

Introduction

The interaction between linguistics and lexicography has been fruitful through the years. This partnership has led to the birth of a theoretical framework or "metalexicography" (Wiegand 1984: 15) within which it is possible to survey "the history of lexicography" and learn from it, establish "a general theory of lexicography", do "research on dictionary use" and critically evaluate dictionaries. In Wiegand's general theory of lexicography, dictionary typology is viewed as "a major component ... since many of the statements in this theory have to be formulated relative to the typology" (Wiegand 1984: 17). This claim is supported by the current emphasis in meta- and practical lexicography on the needs of the dictionary user.

The guiding principle of user-friendliness dictates that a lexicographer should have a set group of target users in mind before working out the theory of organisation and accompanying typological profile of a dictionary. According to Al-Kasimi (1977: 18) dictionary typology is inseparable from the purpose of a dictionary which in turn is determined by "the needs of the target user" (Gouws 1989: 66). Sadly, the very close relationship between linguistics and lexicography has had a negative, user-unfriendly effect in the current meta-lexicography. The nonlinguistic dictionary and especially the encyclopaedia — a valuable, widely-used reference work — has been marginalised as a result of the linguistic bias. This article aims to partially correct this imbalance.

There has, however, been a call for a more representative typology from many sources. Even though the nonlinguistic dictionary has not received much meta-lexicographical attention, it has been included in dictionary typologies as an opposite to linguistic dictionaries. The encyclopaedia is seen as the prototypical nonlinguistic dictionary; in fact various meta-lexicographers view the distinction between the encyclopaedia and the linguistic dictionary as one of the most elementary typological distinctions (see Gouws 1989: 60, Landau 1984: 5 and Zgusta 1971: 198). Svensén (1993: 2) locates the difference between these types as inherent in the nature of the information they present: "an encyclopaedia communicates knowledge about the world, whereas the (linguistic) dic-

tionary gives information about the individual units in the communication system (i.e. the language)". Gouws (1989: 60) describes the difference in terms of the motivation for the compilation of the reference tool. The encyclopaedia must serve broader cultural needs, whereas a linguistic dictionary is constructed with linguistic purposes in mind.

There are not necessarily fixed boundaries between these two. Encyclopaedic information is for example often found in general monolingual dictionaries. The cognitive school's emphasis on the role of pragmatic information in the construction of meaning indicates a further bond between what is traditionally seen as linguistic information and encyclopaedic information. Contextual information that does not form part of a strictly semantic analysis of a given lexical item, does however have semantic value within the broader framework of meaning a speaker constructs for him-/herself. The encyclopaedia and the linguistic dictionary should then be seen as two points on a scale with nondiscrete boundaries. Much more effort should also be made to explore the lexicographical relevance and value of the encyclopaedia.

Encyclopaedias

Besides this view on the construction of meaning there are also other reasons why more attention should be paid to encyclopaedias in metalexicographical circles. Gouws (1989: 61) states: "The accommodation of encyclopaedias in the dictionary family follows from the principles of arrangement (of entries etc.) that are employed in encyclopaedias and that are motivated by lexicographic criteria" (my translation). More will be made of this in the discussion of guide structures in *Microsoft Encarta 99 Encyclopedia* (henceforth *Encarta 99*).

Encyclopaedias also warrant more attention for a pragmatic reason. Next to general dictionaries, general family reference encyclopaedias are probably the most widely-used, influential reference sources, particularly for younger users. Technological advances have made encyclopaedias even more prominent reference sources. With the advent of the computer age, encyclopaedias are no longer solely available as overpriced, oversized, cumbersome multi-volume books. CD-ROM encyclopaedias are generally not larger than CD-ROM dictionaries and are not more expensive. For those who can afford a computer, these encyclopaedias can be powerful resources of valuable information. A school pupil, for example, will not have to go through the torturous experience of library visits, difficult search routes through several volumes and often more than one copying and/or photocopying session. All the information he/she requires (including pictures) is available at home on computer. These developments do not however render their predecessors obsolete, and there will probably still be a market for printed dictionaries and encyclopaedias for a very long time.

General family reference encyclopaedias on CD-ROM

The commercial sector has long since seen the potential of computerised innovation. Several encyclopaedias are currently competing on the computer market, including products made by Grolier and Compton from the United States of America and encyclopaedias by Hutchinson and Oxford in Great Britain. Yet one of the most successful family reference works is brought to the market by Microsoft, the creators of the Windows interface and operating system.

Most of these encyclopaedias, including Microsoft Encarta, have been available for some time, and with a new, updated version appearing every year the compilers have taken care to maintain the pace and gain any possible advantage over competitors. The result is that Encarta 99 is a formidable opponent combining the latest technology with a solid, well-researched informative text as part of the microstructure.

The microstructure

In this article the central focus is on the ways in which Encarta 99 guides the user to the required information and not on the information itself. The microstructure does however reveal a relevant aspect of this encyclopaedia: who its target users are. This task should be performed by the introduction in the *help* (or *overview*) files and/or by an opening statement in the additional instruction booklet, but sadly it is not done. Should one then not assume that a general family reference encyclopaedia should meet the needs of the whole family? This is certainly the ideal, but does not reflect reality. The possible age span of a family and the specific requirements of each family member make it virtually impossible to meet the needs of for example the preschool, primary school, secondary school, tertiary student and professional adult member of the family in one CD-ROM encyclopaedia.

Encyclopaedias do tend to focus on specific target groups. The complexity and information density of the microstructure of for example *Compton's New Century Reference II* (henceforth abbreviated as *Compton's*) seem to suggest tertiary students as its focus group, whereas the 1997 *Grolier Multimedia Encyclopedia* (henceforth *Grolier*) probably aims primarily at secondary school pupils. Encarta 99's text is the most versatile of the three, making it accessible to target groups from late primary school level up. *PC Pro* (1997: 215) comments on this element in Encarta 98:

Many articles have been restructured and rewritten to make them more appropriate to either primary or secondary school use. Look up "castle", for example, and you'll find an article written in relatively straightforward language, appropriate to a typical primary school project. Go for something more complex like "quantum theory", and all the terminology

and ideas are presented using advanced language appropriate to secondary school or college users.

Encarta also employs an easily accessible and understandable dictionary that can help to explain difficult words in the microstructure. One might then expect the same versatility to be employed in the "guide structures" (see Louw 1998: 97) of this encyclopaedia. Generally this is so, but it is not always the case.

Guide structures in Encarta 99

In keeping with the broad theme of user-friendliness it is apt that a discussion of a computerised nonlinguistic dictionary should focus on those elements that have revolutionised the way lexicographers think about dictionary accessibility. In a written dictionary or encyclopaedia the search paths users follow to the information they require are almost always relatively slow. Louw (1998: 97-144) shows that lexicographers' creative use of guide structures such as the access, addressing and mediostructures, and functional help texts such as indexes, user's guides, abbreviation lists, etc., can speed up the searches. Yet the restrictions of the book format limits the ingenuity of the lexicographer much more than the freer approach allowed by the computer.

Two of the guide structures mentioned above, the access structure and mediostructure (the structure encompassing all lexicographic cross-references), will be used in the rest of this article to illustrate the tremendous potential of the CD-ROM encyclopaedia and the successes and shortcomings of Encarta 99.

CD-ROM encyclopaedias seem to be the state of the art in accessibility and linking of material. This is probably the result of the urge to continuously improve the product and to keep one step ahead of the competition. Therefore the scope of the terminology used for printed dictionaries will have to be broadened considerably. For example, the mediostructure has to accommodate the hyperlinks between CD-ROM and the Internet facilitated by the hard drive. This development has significantly altered the concepts of intertextuality and cross-referencing (more about this later). Yet the existing terminology still provides a valuable framework within which to discuss Encarta 99. Consider for example the theory of the access structure.

Access structures in Encarta 99

The route to the information desired can still be subdivided into an outer search path and an inner search path. The former takes the user to the desired article while the latter refers to the search conducted in the microstructure to find relevant pieces of information. The structures that are set up to guide the users on these paths by simplifying access to the necessary information are the outer and inner access structures.

The outer access structure

It is the outer access structure that provides the first proof of the CD-ROM encyclopaedia's potential. The rapid outer access structure which encompasses those elements that guides one up to the lemma, needs to be very well constructed. Successful, easy access is imperative in this very early part of the user's search for the required information. The opening screen of Encarta 99 already provides the user with immediate access to the encyclopaedia articles as main macrostructure or direct links to the accompanying dictionary, on-line updating facilities, *help* (or *overview*) files or to the inserted media texts (which can be accessed either as inserted inner texts in the encyclopaedia articles or in this case as independent lists organized into different categories). Encarta 99's opening screen is an attractive, well-designed first stop on the rapid outer search path.

If one chooses to go directly to the encyclopaedia articles there is yet another excellent element of the rapid outer access structure to guide one, viz. the *pinpointer*. This floating help text acts as an index in which all the encyclopaedia articles are presented. One can simply browse through the list using one's mouse and click on the relevant article to display it on the main part of the screen, or even easier, type in the required article name and press the enter key to automatically display the article. If the topic one is searching for is not worded as in the entries in the alphabetical list, there is also a word search feature in the *pinpointer*. The results of simple word searches are however often haphazardly arranged and one might have to wade through irrelevant material before arriving at the required article. The *help* files in Encarta does, however, give the user a detailed account on how to use the different and sometimes complicated search methods. To some extent this makes up for the lack of organization in the results of simple searches, because guidelines are given to help the user refine the search and go directly to the relevant information. With the current emphasis on an immediate transfer of information, the question remains whether a family research encyclopaedia should assume that its users will study the user's guide or *help* files before starting their searches. The creators of Encarta can learn much from the organized, efficient searches allowed in Compton's *idea search* subsection.

The two encyclopaedias mentioned above and Grolier might differ on the format of especially the rapid outer access structure, but with the standard outer access structure (which corresponds to the main macrostructure in each encyclopaedia) the procedures followed are very similar. They all, for example, employ an initial alphabetical listing of articles rather than a thematic listing as is used by the creators of the electronic version of the *American Concise Encyclopaedia*. For the more crafty researchers Encarta 99 offers a secondary thematically structured search path. One can enter this search path by clicking on the button marked *categories* on the right hand side of the *pinpointer*. A list of broad thematic groupings will appear to the right, which can be refined yet again

before a final list of results is given in the main window of the *pinpointer* in alphabetical order. This versatility in its outer access structure shows Encarta 99's commitment to meeting the needs of its broad group of target users. Yet it is the preference for an initial alphabetic listing as the first step in searches using the *pinpointer*, that proves this commitment to the needs of target users, as an alphabetical listing would probably be preferred by families within the current dictionary culture. The adherence here to lexicographic criteria corresponds with Gouws's (1989: 61) motivation for including encyclopaedias in the dictionary family.

The inner access structure

The principle of user-friendliness also guided the compilers in their creation of the inner access structure. The most innovative part of Encarta 99's inner access structure is the *outline* given with subdivided articles. If one looks up "dictionary" one would find an *outline* containing a general introduction and four sub-headings "Encyclopedic Dictionaries", "Earliest Dictionaries", "Major Dictionaries of Continental Europe" and "Dictionaries in the English Language". By engaging the *outline + text* option at the *view* function the user can ensure that the *outline* will be present at the left side of the screen for as long as he/she wants it there. In articles that are not subdivided that portion of the screen will remain empty. A most welcome addition is that a click on the desired subsection in the *outline* will move the main text to that point as well. This is useful in very long articles such as the one on World War II.

Another *view* option is to have the relevant media elements in smaller size on the screen at the same time as the *outline* and the main text. This may however lead to a very crowded display that defies the principle of user-friendliness. The *outline* can therefore be removed, leaving only the media element and main text. One can of course exercise the option of having only the main text or the media element on screen, if so desired.

Another important aspect of the inner access structure, is the easy access to the media elements either from the main text (where they operate as hidden inserted inner texts easily accessible by means of a double click on a relevant icon), or from the *outline*. The availability of media features in the *outline* can save a prolonged search through the main text and is one of Encarta 99's unique features.

Yet it is this same interaction between media elements, *outline* and main text that leads to one of the user's greatest frustrations. The inner search path is often rudely interrupted when one is calling up media elements by a prompt to change the CDs. Owing to Encarta's size it had to be accommodated on two CD-ROMs. The *Deluxe* edition adds a third with an enhanced search engine called a research organiser which is "great for school projects" (Hardaker 1997: 23) and teaches students the research skills they need to deal with vast amounts of information. The changing is inevitable, but a more systematic approach as

to which media elements are on which CD would be welcome. At the moment certain media elements (even of the same type) in a given article are on one CD and the rest on the other. The confusion is compounded by the fact that no explanation or policy statement regarding the distribution of material on each CD is given in the *help* files. This situation should be remedied before new editions of Encarta appear.

The mediostructure

Innovation in lexicographic practice is not limited to the outer and inner access structures. Encarta 99 also takes the art of lexicographical cross-referencing to new heights. Certain words in the main text are displayed in red and a click on one of these transports the user to another related article. In the article with "dictionary" as lemma, one would find in the subsection "Dictionaries in the English Language" the name "Noah Webster" in red. A click on these words takes the reader to an article on his life's work. A click on the icon denoting *back* then guides the user back to the previous article "dictionary", if he/she wishes to go there. Encarta 99's cross-references are not only clearer than e.g. Compton's or Grolier, but "compared with the other two it also has more hypertext links within articles which take you to other, relevant articles" (Hardaker 1996: 46).

Unlike passive cross-referencing in printed dictionaries these active cross-references provide the user with direct access to the required information. This increases the speed of the search, saving the user time and effort. The innovation does not stop with these cross-references to external articles though. Encarta 99 expands the bounds of intertextuality by employing active text-external cross-references which would have been impossible in books. Firstly, one can click on almost any word in the main text of an article that is not a name and be greeted by a concise definition courtesy of Encarta's borrowed, integrated dictionary (in this World English Edition of Encarta 99 the 9th edition of the *Concise Oxford Dictionary*).

Secondly, several articles (see for example "World War I") include an icon that if clicked on calls up a relevant segment of text in the broader text called the *timelines*. These segments are not inserted texts, because the whole *timelines* text is active at the time it is displayed on screen (which is not necessarily the case when a map is inserted from the atlas). The *timelines* link is therefore an effective active article-external text-external cross-reference.

As active encyclopaedia-external cross-references, Encarta 99 offers (as does most of its competitors) links to the Internet. Hardaker (1996: 42) states about these encyclopaedias: "In their latest incarnations they connect to on-line services and thus automatically update themselves". Laborious supplements are replaced by swift, monthly updates to ensure that the user stays in touch with the ever-changing reality.

The more internal references (especially the article-internal references) are not to be outdone, however, by those covering greater distances in the search path. As has already been shown, the active cross-references in the *outline* can save the users time in their inner search path and are mostly direct and relevant. The exception is once again active cross-references (especially to multimedia inserted texts) that lead to a prompt for a change of CDs, but even with this minor delay the user still gets to the required information faster than with any printed encyclopaedia.

Active cross-references, as employed in Encarta 99, should also form part of a discussion of the future of linguistic dictionaries. It is one of many new ideas that make general linguistic dictionaries on CD-ROM or the Internet exciting new tools especially for those who work with a word processor. Whereas monolingual descriptive English dictionaries are already well represented on CD-ROM (with the first Afrikaans standard descriptive CD-ROM dictionary recently having appeared), translation dictionaries are not readily available yet. There is for example a dire need for such a dictionary with Afrikaans and English as treated language pair, which can be consulted when using a word processor.

Other helpful additions

In addition to the revolution in guide structures such as the access and mediostructure, CD-ROM encyclopaedias have also brought a new dimension to the main text, either to help the user to understand the program or decode the main text or to help the user by providing additional information to the main text. Multimedia has provided these texts with the benefit of sound and in some cases video. Probably the best illustration of the added value is provided by those media elements acting as inserted inner texts.

Inserted inner texts

These easily accessible texts come complete with captions that often summarize relevant parts of the main text article. The inserted texts can act independently, but are at best employed in co-operation with the main text. If one goes to the article on World War II one finds a vast array of media elements, including several photographs, audio texts (e.g. on the signing of the Munich Pact), tables, etc. The *Deluxe* edition even includes *sidebars*, inserted texts that are easily accessible from both the main text and the *outline* and "include personal essays, topical news stories, case histories, or other interesting and unique viewpoints on the topic you're exploring" (Encarta 99 *Deluxe* User's Guide, 17). In the above-mentioned article one finds several of these texts that provide a unique, undiluted view of the war. A student can for example access King

George VI's broadcast on the eve of war, Roosevelt's Pearl Harbour speech or the English text of the German surrender at Reims.

In some cases these inserted media elements form part of other texts. If one calls up "South Africa" in a separate search for example, a map of South Africa is available as an inserted text. This map is part of a complete atlas which is included on the CD and can be accessed as a whole from the *media gallery*. These inserted texts together with the integration of other texts such as the dictionary, *interactivities* and *timelines*, as well as the *virtual tours* of the *Deluxe* edition, prove that the encyclopaedia is not a text, but a complex text type carrier which allows free movement between its constituent texts. Though this is often also the goal of the printed (linguistic or nonlinguistic) dictionary the freedom of access granted by the computerized format means that a CD-ROM encyclopaedia such as Encarta 99 comes much closer to reaching this goal.

Towards the future

From this article the tremendous advantages offered by CD-ROM encyclopaedias should be apparent and yet they are already under threat of being surpassed by the Internet. At the moment computer users in search of nonlinguistic dictionaries make do with CD-ROM encyclopaedias or encyclopaedia downloads from the World Wide Web. With the improvements being made on the search engines of the Internet to make search paths faster, together with the availability of books, audio and video texts on the Internet, it is conceivable that the Internet can function as an almost limitless encyclopaedia. With its tremendous speed and information storage capacity, this "encyclopaedia" can dwarf printed and CD-ROM versions and redefine our concept of traditional lexicography. It also offers practical lexicography and metalexicographers a world of new opportunities if they can detach themselves from a linguistic bias as well as conquer the technology of the computer.

Bibliography

Encyclopaedias

American Concise Encyclopedia on Power CD. 1995. ZCI Publishing Inc.

Compton's New Century Encyclopedia and Reference Collection II. 1995. Compton's New Media Inc.

1997 Grolier Multimedia Encyclopedia. 1997. Grolier Interactive Inc.

Microsoft Encarta 99 Encyclopedia. 1998. Microsoft Corporation.

Other sources

- Al-Kasimi, A.-M. 1977. *Linguistics and Bilingual Dictionaries*. Leyden: E.J. Brill.
- Gouws, R.H. 1989. *Leksikografie*. Cape Town: Academica.
- Hardaker, M. 1996. CD-ROM Encyclopaedias. S.A. *Computer Magazine*, September 1996: 40-48.
- Hardaker, M. 1997. Encarta 98. @ *home*, November 1997: 23.
- Hartmann, R.R.K. (Ed.). 1984. *LEXeter '83 Proceedings*. Tübingen: Max Niemeyer.
- Landau, S. 1989. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Louw, P.A. 1998. *Kriteria vir 'n standaard vertalende woordeboek*. Unpublished M.A. thesis. Stellenbosch: University of Stellenbosch.
- PC Pro*, Issue 38, December 1997: 215.
- Svensén, B. 1993. *Practical Lexicography*. Oxford: Oxford University Press.
- Wiegand, Herbert E. 1984. On the Structure and Contents of a General Theory of Lexicography. Hartmann, R.R.K. (Ed.). 1984: 13-30.
- Zgusta, I. 1971. *Manual of Lexicography*. The Hague: Mouton.