# Lexicography in the Information Age[*]

Sven Tarp, *Centre for Lexicography, Aarhus School of Business, Aarhus, Denmark (st@asb.dk)*

**Abstract:** Lexicography has much to contribute to the information and knowledge society. The very essence of lexicography is its capacity to provide quick and easy access to data from which information needed by different types of users in different types of social situations can be retrieved. In this respect, the article discusses the concepts of information and knowledge related to lexicographic theory. On this basis, it argues that lexicography, if it concentrates on its core capacity, has the potential to be projected beyond the limits of known dictionaries. It proposes that besides the communicative and cognitive user situations that have already been dealt with in lexicographic literature, a third situation, the so-called operational situation, should also be added to the situations relevant to lexicographic theory. In this regard, it points out that other reference works than those traditionally discussed in lexicographic theory should be added to the list of utility tools that may benefit from a lexicographic approach.

**Keywords:** LEXICOGRAPHY, LEXICOGRAPHIC TOOLS, INFORMATION NEEDS, INFORMATION AND KNOWLEDGE SOCIETY

**Opsomming: Leksikografie in die inligtingsera.** Leksikografie het baie om by te dra tot die inligtings- en kennisgemeenskap. Die werklike essensie van die leksikografie is sy vermoë om vinnige en maklike toegang te verleen tot gegewens waarvan inligting benodig deur verskillende soorte gebruikers in verskillende soorte maatskaplike situasies verkry kan word. In hierdie verband bespreek die artikel die begrippe inligting en kennis met betrekking tot die leksikografiese teorie. Op grond hiervan voer dit aan dat die leksikografie, indien dit op sy kernvermoë konsentreer, die moontlikheid besit om verby die grense van bekende woordeboeke gevoer te word. Dit stel voor dat, benewens die kommunikatiewe en kognitiewe gebruikersituasies wat reeds in die leksikografiese literatuur behandel is, 'n derde situasie, die sogenaamde operasionele situasie, by die situasies tersaaklik vir die leksikografiese teorie gevoeg behoort te word. In hierdie verband toon dit aan dat ander naslaanwerke as dié wat tradisioneel in die leksikografiese teorie bespreek word, ook by die lys nutsgereedskap wat uit 'n leksikografiese benadering mag voordeel trek, gevoeg behoort te word.

**Sleutelwoorde:** LEKSIKOGRAFIE, LEKSIKOGRAFIESE GEREEDSKAP, INLIGTINGSBEHOEFTES, INLIGTINGS- EN KENNISGEMEENSKAP

---

[*] This article was presented as a paper at the Twelfth International Conference of the African Association for Lexicography, organised by the Tshwane University of Technology in collaboration with the IsiNdebele National Lexicography Unit, at the Soshanguve Campus, Pretoria, Republic of South Africa, 27–29 June 2007.

It is frequently stated that we are living in a so-called information and knowledge society. Such statements are often highly superficial, for example when this kind of society is considered as synonymous with a post-industrial society where industry allegedly should have given way to something else. With this claim, industry is mainly defined according to some historically ephemeral features, i.e. its temporary fordist features, and not according to its more general role as the most important organisational form of human production even today. But apart from such examples where secondary features are transformed into primary ones, it cannot be ignored that knowledge and information have become increasingly important components of present-day society in general and post-fordist industry in particular. In this respect, it is acceptable to speak of an information and knowledge society.

But what has all this to do with lexicography?

When an abstraction is made from the dictionary form, i.e. from both the traditional printed dictionary and the more advanced electronic one, what is left is a utility tool whose essential characteristic is its capacity to provide quickly accessible data from which information can be retrieved. And these data are furthermore selected and prepared in order to satisfy the information needs, not of a specific type of users in general, but of a specific type of users finding themselves in a specific type of extra-lexicographic situation where such needs may emerge. This is the very essence of lexicography which can be projected far beyond the traditional limits of known dictionaries. This contribution will analyse this idea and outline some of its interesting perspectives.

## Lexicographic needs

Although Wiegand (1988) was the first to state it explicitly, the idea that dictionaries are utility tools which should be designed for their users is not new. For instance, at the end of a classic conference on lexicography in 1960, Householder (1967: 279) made an important recommendation that has been quoted repeatedly, not least in English-language lexicography:

> Dictionaries should be designed with a special set of users in mind and for their specific needs.

At an abstract level, it is hard to disagree with this recommendation. But at least two fundamental factors are missing in terms of the potential users' needs which in a certain way are left in a vacuum. In the first place, users do not have specific needs unless they are related to specific types of situations. Consequently, it is not enough to define which types of users have which needs, but the types of social situations in which these needs may arise should also be described. Not all social situations are relevant to lexicography, but only situations in which needs may arise that can be satisfied by consulting dictionaries or other lexicographic tools.

Secondly, even in these relevant situations, not all needs are lexicographi-

cally relevant. If a translator, for example, becomes hungry while practising his/her profession, this need for food is completely irrelevant to lexicography. Hence, the general type of needs that are lexicographically relevant should also be defined and later subdivided according to the various types of relevant user situations and user profiles.

As it has already been mentioned, the only types of needs that can be met by dictionaries and other lexicographic tools are the needs for information. Users consult dictionaries in order to obtain information which they can then use for a great variety of purposes, e.g. to solve communication problems or to add to their existing knowledge. However, dictionaries themselves do no contain information, but only lexicographically selected and prepared data from which the users may or may not be able to retrieve the needed information. Wiegand (1998, 2000, 2002) introduced this important distinction between what is contained in the dictionaries and what is retrieved from them through a cognitive process during a successful consultation. His distinction is based on a philosophical approach according to which knowledge — and information — cannot exist separated from the human brain.

It could be argued that such a philosophical discussion is irrelevant to lexicography and that lexicographers would surely have different opinions on a crucial question of this kind. However, even if some lexicographers disagree with Wiegand's basic premises due to their extra-lexicographic world view, it would be very useful to maintain the distinction between the lexicographic data contained in the dictionary and the information extracted from these data. If a dictionary user, for instance, wants to know something about geology and consults a Chinese encyclopaedic dictionary of geology, the information that can be retrieved depends on the user's proficiency level in Chinese and not on the correctness of the lexicographic data. Although these data are not prepared for non Chinese-speaking users, who will probably not be able to extract any useful information at all, this does not mean that it is impossible for the anticipated type of users to retrieve the needed information. Besides, such specialised dictionaries of geology should not only be conceived for speakers of a specific language, but they should also be adapted to the users' scientific level, i.e. whether they are experts, semi-experts or lay people. A lay person, for instance, will probably not be able to retrieve any useful information from a dictionary designed for an expert although all the necessary data are included in the dictionary, but in such a complex way that these are almost incomprehensible without a specialised knowledge of geology. Thus, lexicographic data is something different from the information retrieved from these data.

### Data and quick accessibility

There are various written and oral sources from which human beings can acquire information for some purpose or other. Among the written sources, the most important are newspapers, journals, magazines, books, text books, inter-

net-based texts and dictionaries. In this respect, no information that cannot be retrieved from these sources is obtained by the consultation of dictionaries. For many purposes, some of these sources, especially text books, may be even more appropriate than dictionaries. Nevertheless, dictionaries have some specific benefits which make them far more suitable for other purposes.

— Firstly, dictionaries furnish lexicographically selected and prepared data which are, at least in theory, especially adapted not only to specific types of users (something which they have in common with other written sources), but also to specific types of user situations (where they generally have few rivals).

— Secondly, and even more importantly, the relevant data included in dictionaries are made easily and quickly accessible through different kinds of prepared and foreseen access routes (cf. Bergenholtz and Gouws 2007, and Tarp 2008).

Accessibility is a key concept in any lexicographic theory maintaining to be user-oriented. However, it could be claimed that any data included in any text is accessible to anyone who takes the time to read the text through from beginning to end. Other tools, especially browsers connected to the internet, also provide easy access to relevant data, but as the search on the internet in most cases produces much redundant material in need of further processing, it can hardly be considered quick.

Hence, one of the really distinctive features of dictionaries and other lexicographic tools is that they provide quick and easy access to the specific types of data that can cover a specific type of user's specific types of information needs in a specific type of extra-lexicographic situation. In this respect, lexicography is highly relevant to a society where information is increasingly important to all its members in their professional and daily life and which even defines itself as an information and knowledge society.

## The relation between information, knowledge and communication

There is a large variety of uses and definitions of the word *information*. In Wikipedia (June 15, 2007) the following short explanation of *information* can be found:

> **Information** is the result of processing, manipulating and organizing data in a way that adds to the knowledge of the receiver. In other words, it is the context in which data is taken. Information as a concept bears a diversity of meanings, from everyday usage to technical settings. Generally speaking, the concept of information is closely related to notions of constraint, communication, control, data, form, instruction, knowledge, meaning, mental stimulus, pattern, perception, and representation. Many people speak about the Information Age as the advent of the Knowledge Age or knowledge society, the information society, and information technologies, and even though informatics, information science and

computer science are often in the spotlight, the word "information" is often used without careful consideration of the various meanings it has acquired.

It is logical that the definition of central concepts frequently varies from science to science, and even from theory to theory within the same science, as they all have their own specific approaches to these concepts. The same applies to lexicography which needs it own definitions, not only of *information*, but also of *knowledge*, *communication* and other important concepts. In this respect, the lexicographic function theory agrees with the above definition that "information is the result of processing […] data", but it does not agree that the information is necessarily something processed, manipulated and organised "in a way that adds to the knowledge of the receiver". The function theory sees information as a kind of "proto-knowledge" with the potential to be transformed into knowledge, but also to be used for other purposes. In this regard, knowledge is considered as information stored in the brain — through cognitive processes which is not the study field of lexicography — for subsequent use.

There are two main reasons for this clear distinction between information and knowledge in the lexicographic approach. Firstly, dictionary usage shows that users often, within a period of some hours or days, consult the same entry two or three times in order to satisfy their need for the same specific information. In such cases, the information originally retrieved from the data contained in the dictionary may have served to meet the users' immediate needs, but it has obviously not been completely and satisfactorily stored in their memory for subsequent use.

Secondly, not all dictionary consultations are aimed at adding to the users' knowledge. For instance, when it is a question of problems or doubts arising during the translation of texts from one language into another, or during other communication processes, the purpose of the dictionary consultation is exclusively to find answers to these problems and doubts. In such types of user situations, the information retrieved from the dictionary is needed as a kind of "trouble-shooter" that may have only an ephemeral existence. Although it cannot be excluded that the information in such cases is eventually also stored in the memory, it is important to clarify that this is not the direct purpose of the dictionary consultation, but rather a sort of fortunate by-product of this consultation.

In this respect, there is a complex dialectic relation between information, knowledge and communication (cf. Tarp 2005). The very process of retrieving information from a dictionary is essentially a delayed communication between the original compiler(s) of the dictionary and the user. The information needed for immediate communication purposes may eventually be stored in the memory as knowledge, and the information retrieved with the direct purpose of adding to the user's existing knowledge may likewise be used for subsequent communication processes. In order not to become lost in this complex relationship, it is important always to focus on the original type of information needs that gave rise to the consultation process, and this basically concerns the type of user situation.

## Cognitive user situations

The most thoroughly treated and analysed user situations in theoretical literature are the communicative ones, e.g. text production, reception, translation, marking, revision and proofreading. This is probably due to the fact that theoretical lexicography has so far been mostly dominated by scholars with a linguistic background. Whatever the reason, it is of vital importance for lexicographic theory to advance towards a similar understanding of the so-called cognitive user situations which have up till now been defined in contrast to the communicative user situations as those where users for one reason or another want or need to acquire knowledge about a specific phenomenon. This want or need to gain new knowledge can arise in countless social situations, for instance:

(1)     while reading: the sudden need to acquire additional encyclopaedic knowledge in order to understand the text;
(2)     while writing: the need to know more about a given topic in order to finish a text;
(3)     during discussions with other people: the need to clarify a specific issue;
(4)     during processes in the subconscious: the sudden desire to examine something;
(5)     during dictionary consultation: the desire to know more about a specific topic;
(6)     during preparation for specialised translation and interpretation tasks: the need to learn more about the subject field in question;
(7)     in relation to a teaching programme or a course of study: the need to know more about a specific subject field;
(8)     …

The options are unlimited. However, as most people will be familiar with at least some of the above-mentioned user situations, it is possible even on the basis of this limited selection to deduce some preliminary conclusions. Inspired by the distinction which Hausmann (1977) made between punctual and global issues in regard to dictionary consultation, a fundamental distinction can be made between *sporadic* and *systematic* user situations. Of the above seven situations, the first five are isolated situations without any clear relation to previous or subsequent situations giving rise to dictionary consultation, while the two last ones are situations where the potential dictionary user is making — or is supposed to make — a systematic effort to gradually acquire knowledge about a specific subject field and where a specific consultation in one or the other way is related to previous or subsequent consultations. This distinction between systematic and sporadic situations has important consequences for the design of a dictionary conceived to satisfy the potential users' needs, because the needs arising from a systematic study of a given subject field will require a more sophisticated data distribution structure with synopsis articles and special sections containing a global introduction to the subject field in question as

well as a highly developed system of cross-references or -links allowing users to navigate in the dictionary in order to make full use of its data.

In addition to the distinction between sporadic and systematic user situations, there is a second important conclusion that can be deduced from the above-mentioned user situations. In the systematic situations, it is evident that the purpose of the dictionary consultation is *to add to the users' existing knowledge* (cf. the quoted explanation from Wikipedia). However, at least in two of the other situations, (2) and (3), the direct purpose of the consultation is *to meet an immediate knowledge need* and whether or not the information is added to the user's existing knowledge, i.e. stored in the memory, is a secondary question for the user. When a journalist, for instance, is writing an article about Napoleon and has to inform the readers about the date of birth of this statesman, he/she may then consult a dictionary with the immediate purpose of finding this specific information which he/she may never need again and may probably forget in a few minutes.

Finally, there is a third interesting conclusion that can be deduced from the seven user situations mentioned. In situations (1), (2) and (6), the needs for knowledge are directly related to a current or future communication process and, as such, they could be included among the needs arising during this process. However, from a lexicographic point of view, i.e. in terms of the possible consequences for dictionary concepts, it is necessary to distinguish between the users' inadequate linguistic and other communicative *skills* on the one hand and the possible lacunae in their general or specialised encyclopaedic *knowledge* on the other.

For instance, if a person is reading the poem "Lullaby of the Onion", written by the Spanish poet Miguel Hernandez and dedicated to his only son, and does not know that it was composed in jail immediately after the Spanish Civil War as a reply to a letter from his wife telling him she was surviving on bread and onions, then this person may find symbolisms in the poem that do not actually exist. In this respect, additional encyclopaedic knowledge is necessary to understand the general context in which the poem was written and, thus, the poem itself. As regards text production, a distinction should be made, in an analogous way, between *what* is to be communicated and *how* it should be communicated where the latter concerns the text producer's linguistic and communicative skills while the former concerns his/her general or specialised encyclopaedic knowledge. The same holds true for translation where translators frequently need knowledge about either the general context or the subject field of the text to be translated. This knowledge is of a different nature from the translators' general translation skills, including their proficiency level in terms of specialised vocabulary.

Those engaged in the three mentioned types of communication (1) may have the necessary knowledge in advance, (2) may acquire it through a conscious study before starting the communication process, and (3) may acquire it when their lack of knowledge becomes problematical during this process. It is

evident that both situations (2) and (3) are part of the communication process in the broad sense of the word. However, if situation (2), i.e. a conscious study of a given subject field, occurred without relation to a communicative situation, it would clearly be considered a cognitive situation. On the other hand, the only real difference between situations (2) and (3) is temporal, i.e. between the different moments when the needs occur which give rise to a lexicographic consultation. In this regard, the two types of situations are, in fact, cognitive situations related to communicative situations and could be called *communication-related cognitive sub-situations*.

Although some lexicographers may consider this discussion too extreme and academic, the theoretical conclusions drawn from it are of great relevance for lexicographic practice in as much as the user needs arising in the different situations require different lexicographic solutions and in many cases also dictionary compilers with completely different skills.

## New user situations

So far, the lexicographic function theory has discussed two fundamental types of user situations, the cognitive ones and the communicative ones. However, in theoretical literature there are some sporadic references to reference works such as handbooks, manuals and how-tos. These references have not been properly discussed in any known contribution, but it is easy to see that many of these reference works, especially the how-tos, are not designed to provide assistance in either cognitive or communicative situations, but to give instructions on how to proceed in specific situations, e.g. in relation to the operation of machines and other instruments. These situations which may be called *operational situations* have so far not been integrated into the function theory or any other lexicographic theory, for which reason the above-mentioned reference works may not be considered lexicographic tools either.

However, although it is evident that not all handbooks, manuals or how-tos are designed as lexicographic products, there is no reason why they should not be conceived and considered as such. They all have something fundamental in common with traditional lexicographic products, i.e. they are tools conceived to be consulted by specific types of users in order to satisfy specific types of information needs in specific types of social situations. In this respect, they are expected to provide quick and easy access to the relevant data from which the needed information can be retrieved.

It is evident that a lexicographic theory focusing on quick and easy access to data from which specifically user- and situation-adapted information can be extracted has much to contribute to the design of a new generation of improved handbooks, manuals, and how-tos. Moreover, to this list of reference works should be added other tools such as user guides which are frequently not only used to be read from beginning to end, but also — and even mainly — for punctual consultation in order to acquire information that can be trans-

formed into instructions on how to proceed in operational situations very similar to those where handbooks, manuals and how-tos may provide assistance. All these reference works traditionally used in relation to operational situations could benefit from a lexicographic approach which analyses the types of information needs that may arise for specific types of users in the various types of operational situations in order to prepare new types of lexicographic tools with quick and easy access to the relevant data.

In this way, the operational situation should be added to the other fundamental user situations included in the lexicographic function theory. Two-legged lexicography has acquired its third leg wherever it is placed.

## Lexicography at a crossroads

After more than 4 000 years of existence, lexicography finds itself at a crossroads. It will either throw itself into the future with all its renewed theoretical and practical strength or it will become circumvented and confined to the sphere of traditional dictionaries considered as products of applied linguistics. There are three reasons for this challenging situation: firstly, the theoretical efforts carried out during the last decades which have revealed the real essence of lexicography as the capacity of providing information; secondly, the advent of the information and knowledge society; and thirdly, the development of the new computer and information technologies.

In order to meet the new challenges, lexicography must, on the one hand, focus on its core knowledge and, on the other, project itself far beyond its traditional limits. It should further study and develop its knowledge about user situations, user profiles, information needs, data selection and presentation, and quick and easy data accessibility. It should promptly and totally adapt itself to the new technologies and it should, among others, explore the possibilities of performing user needs adapted searches on the internet combining traditional static data with dynamic data made available through the internet.

There is no doubt that due to its focus on information needs related to specific types of users and user situations, lexicography is in a position to furnish essential contributions to the new information sciences. In this respect, it may be considered whether lexicography should not change its heavily connotative name and find a new name which corresponds far more to its real essence. Perhaps *infology* or *informology* may qualify as candidates in this search for lexicography's present-day and future identity.

## Literature

**Bergenholtz, Henning and Rufus H. Gouws.** 2007. The Access Process in Dictionaries for Fixed Expressions. *Lexicographica* 23. [Forthcoming.]

**Hausmann, Franz Josef.** 1977. *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen: Niemeyer.

**Householder, Fred W.** 1967. Summary Report. Householder, F.W. and S. Saporta (Eds.). 1967. *Problems in Lexicography*: 279-282. Bloomington: Indiana University.

**Tarp, Sven.** 2005. The Pedagogical Dimension of the Well-Conceived Specialised Dictionary. *Ibérica. Journal of the European Association of Languages for Specific Purposes* 10: 7-21.

**Tarp, Sven.** 2008. *Lexicography in the Borderland between Knowledge and Non-Knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Lexicographica. Series Maior. Tübingen: Max Niemeyer. [Forthcoming.]

**Wiegand, Herbert Ernst.** 1988. Was eigentlich ist Fachlexikographie? Mit Hinweisen zum Verhältnis von sprachlichem und enzyklopädischem Wissen. Munske, H.H. et al. (Eds.). 1988. *Deutscher Wortschatz. Lexikologische Studien. Ludwig Erich Schmitt zum 80. Geburtstag von seinen Marburger Schülern*: 729-790. Berlin/New York: Walter de Gruyter.

**Wiegand, Herbert Ernst.** 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie. 1. Teilband.* Berlin/New York: De Gruyter.

**Wiegand, Herbert Ernst.** 2000. Wissen, Wissenrepräsentationen und Printwörterbücher. Ulrich, H., S. Evert, E. Lehmann and C. Rohrer (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000. Stuttgart, Germany, 8–12 August 2000*: 15-38. Stuttgart: Universität Stuttgart, Institut für Maschinelle Sprachverarbeitung.

**Wiegand, Herbert Ernst.** 2002. Wissen in der Sprachlexikografie. Ein Plädoyer für einige immer noch notwendige Differenzierungen. Ezawa, K., W. Kürschner, K.H. Rensch and M. Ringmacher (Eds.). 2002. *Linguistik jenseits des Strukturalismus. Akten des II. Ost-West-Kolloquiums Berlin 1998*: 265-281. Tübingen: Gunter Narr Verlag.

*Wikipedia. The Free Encyclopedia.* http://en.wikipedia.org.