Sven Tarp*

Revival of a Dusty Old Profession

Abstract
This text is the inaugural lecture presented by Professor Sven Tarp at the Aarhus School of Business on March 14, 2008. Firstly, the text provides a brief retrospect of the history of lexicography with emphasis on the experience of the big Chinese encyclopedias and the first big National Danish Dictionary. On this basis, it calls for the further development of an independent theory of lexicography in order to go beyond the experience of past and present lexicographic works and project the discipline into the future. It then discusses some of the problems hampering this process. With a call for innovation, it urges the lexicographers to produce the works that are really needed by users and the State to finance the production of such tools of national, cultural, social and economic importance in the present information era. Finally, it concludes that the researchers in lexicography need the audacity to go beyond the usual boundaries and generate new ideas, even if they are initially not welcomed or even understood.

Over the years, many colleagues, friends and family members have asked me to explain what I am actually doing and what is really meant by *lexicography*. I must admit that I myself as well as other researchers within this field of knowledge have asked the same question many times, although for other reasons. The term *lexicography* is far from clear and unambiguous. One way to pinpoint its meaning is to go back in history and determine the origin of the term, which can be found in the Greek words *léksis* that means “speech, word”, *leksikón* which is a book “dealing with words” and *gráphein* that means “to write”. In this light, lexicography would mean something like “the writing about words”, i.e. a very broad definition which actually embraces any kind of text related to linguistics and dealing with language.

* Sven Tarp, Professor, Dr.ling.merc.
Centre for Lexicography
Aarhus School of Business, University of Aarhus
Fuglesangs Allé 4
DK-8210 Aarhus V
stt@asb.dk

Hermes – Journal of Language and Communication Studies no 41-2008
Harmless drudgery

However, at a certain moment in history, the terms *lexicography* and *lexicographer* became specifically related to the writing and writer of dictionaries. We don’t know exactly when the two terms were first used in this sense, but there is no doubt that it happened many centuries ago as can be seen in the following dictionary article, published in 1755 in the *Dictionary of the English Language* and written by a harmless drudge who went down in history as one of the most famous lexicographers of all times:

**Lexicographer**
A writer of dictionaries; a harmless drudge, that busies himself in tracing the original, and detailing the signification of words.

It may not be fair to consider Samuel Johnson, the person who busied himself in writing this dictionary article, as “dusty” in the light of the humour and prejudice which he, unlike most other lexicographers, introduced in his definitions but in one way or another he could rightly be considered to be “old”. However, the definition of a lexicographer provided by Samuel Johnson in the above article is far from adequate. It goes without saying that lexicographers do “write” dictionaries but they also write other types of similar “books”.

In fact, only four years before the publication of Johnson’s dictionary, on the other side of the English Channel, the first volume of the world famous French *Encyclopédie* was published with the exact title: *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers* (Encyclopedia, or a systematic dictionary of the sciences, arts and crafts). Here we see that the editors had decided to include both the word *encyclopedia* and the word *dictionary* in the very title of the book, and the reason for this was not that they did not know the significance of the two words because, among the total of 71,818 articles included in the 35 volumes, there is one about encyclopedias and another about dictionaries, written by each of the two main editors, Diderot and d’Alembert, respectively. Historically in fact, lexicography does not only comprise dictionaries, but also encyclopedias, lexica, thesauri, glossaries and a number of other types of reference works. Recently, theoretical lexicography has even started to deal with how-to’s, handbooks, manuals, etc.
The truth is that the present-day meaning of the term *lexicography* cannot be traced by means of etymology or determined through definitions given by harmless drudges at different moments in history, but only on the basis of an up-to-date knowledge of the theoretical and practical development of this branch of human activity. In this respect, the meaning of the word *lexicography* is continuously changing, and it should be of no surprise to anybody that today lexicography is moving the gun into position in order to conquer new land and occupy a prominent place in the so-called information and knowledge age.

Both the French *Encyclopédie* and Johnson’s *Dictionary of the English Language* may be considered old lexicographic works. But in a historical perspective their 250 years make them relatively young compared to the first known dictionaries which were compiled more than 4,000 years ago in India and later in China and Mesopotamia. Since then, hundreds of thousands if not millions of different dictionaries and other lexicographic works have been published, and if we count the single copies, the number runs into billions.

These lexicographic works are carved in clay, hand-written on paper or papyrus, type-written, printed with different technologies or made available electronically on compact disks and internet. They are written in hundreds of different languages, and even in dialects. They cover almost all spheres of human activity and knowledge. They have a huge variety of purposes. They transmit knowledge between generations and within each generation. They help people to communicate in their mother tongue and between the language communities. They assist their users with advice and instructions in order to perform manual or mental actions. In short, dictionaries and other lexicographic works have played and still play an extremely important, although frequently neglected, role in our culture and social development.

The cradle of encyclopedias

There are small lexicographic works and big works. Although we Europeans are rightly proud of our big national dictionaries and encyclopedias, many of which comprise dozens of volumes, they cannot stand comparison with the ones produced in other parts of the world, especially in China which can be considered the cradle of encyclopedias.
The first Chinese encyclopedias were written in the Western Zhou Dynasty (1050–771 BC) and have long ago been transformed into dust. The largest and most complete encyclopedia that has survived from imperial China is the *Ch’in ting ku chin t’u shu chi ch‘êng* (Synthesis of books and illustrations of ancient and modern times). It was printed in 1726 in about 60 copies and contains 100 million Chinese characters, 800,000 pages and 10,000 chapters collected in 5,020 volumes. If we take into account that each page with Chinese characters corresponds to about six pages written with our alphabet, it is, as such, the biggest lexicographic work ever printed. But it is not the biggest one ever produced.

The one lexicographic work entitled to this honour is the *Yung-lo ta-tien* (Great Canon of the Yung-lo Era) which was completed in 1408 and comprises a total of 22,937 chapters bound in 11,095 volumes. This great encyclopedia is in many ways a remarkable work from which a lot can be learned even today. It was structured according to a rhyming system for the characters, to which further data were addressed and partially transcribed character for character as exact copies of original texts produced during the previous decades. In this way, the *Yung-lo ta-tien* was the major collective literary enterprise of the Ming Dynasty (1368–1644). The idea was to make a complete Canon of existing texts within a wide range of subject matter, such as the classics, history, philosophy, Buddhism, Taoism, drama, arts, and farming, at a critical historical moment when China was recovering from several devastating conflicts and needed this knowledge.

The proposal for the undertaking was officially made by the Emperor’s Grand Secretary on July 19, 1403. Only 17 months later, on December 13, 1404, the first draft was presented to the Emperor who was not satisfied at all and therefore decided to enlarge the staff from 141 to 2,169 assistants in order to finish the much desired work. On December 17, 1408 – i.e. in an admirably short span of only five years – the manuscript which occupied 40 cubic metres was finally completed and ready for publication. However, the mighty Chinese Emperor then discovered that even he did not have enough money to print it with the technology of that time. For this reason, only two hand-written copies were made and today most of the material has been lost due to fires, dynastic wars and foreign invasions.
There are at least two important lessons to be learned from this gigantic enterprise. The first is that it is necessary for the State to intervene in order to finance such big projects which are needed in order to enhance the social and cultural development of a nation. If this is not the case, a number of problems may occur. In the title of this inaugural lecture I refer to lexicography as a “profession”, but this should be taken with a grain of salt. Most authors of dictionaries are not professionals and do their lexicographic work in their spare time, frequently without even getting paid when the work is eventually finished. This may affect the quality of their work, and even worse, it may also imply that a number of much needed lexicographic works are never published because their production depends on the private initiative and is rarely “a good business” in the narrow economic sense of the word.

The second lesson to be drawn from the concrete Chinese experience is that the lexicographers’ knowledge and workforce should be used rationally and the projects carried out and finished within the shortest possible span of years. It is a fact that many lexicographic projects drag out for decades, sometimes even for more than hundred years, and that others may never be finished. This represents a gigantic waste of human energy and frequently implies that the first volume of a big work is more or less outdated and old-fashioned when the last volume is eventually published, often according to a more or less modified concept.

Up till now I hope to have shown that lexicography, though it may rightly be considered an old and dusty profession, has a rich and glorious past. However, the chair at the Centre for Lexicography has not been established in order to study the past, but to develop theories that can guide the present and the future although this, of course, only can be done based upon the experiences of the past. Thus, the first question to be asked is whether or not lexicography has a future. It will be of no surprise that my answer is: yes, it has. I am convinced that lexicography will have a future that will be even more glorious than its past. But it will be a future that in many ways differs from what has been known and done until now. And this difference will inevitably be the product of the combined theoretical and practical creativeness of a lot of harmless drudges engaged in lexicography in one way or another. As I see it there are four major obstacles or problems which hamper and obstruct the way towards the development and improvement of the lexicograph-
ic tools of the future. These four problems are: academic pride, the Market, faster horses and Copenhagen.

**Academic pride**

During the centuries, an artificial and extremely harmful contradiction has been created between academic knowledge and popular wisdom. On this basis, a number of university students and graduates have developed the ridiculous and arrogant habit of boasting of their knowledge and trying to impose it on the ordinary people completely disregarding the wisdom and real needs of the latter. In Denmark we have Holberg’s *Erasmus Montanus*, in Germany they have Lessing’s *Junge Gelehrte*, and in other countries they have other big-headed personalities who have caused people to laugh over and over again.

Unfortunately, these personalities are not limited to literature, but are represented in real life. This became particularly clear when the first big Danish National Dictionary was discussed and planned in the 18th century. The Board of the then Society of Sciences who was supposed to be responsible for the practical implementation of the project came up with a number of Erasmus Montanus-like ideas, among them that the Danish headwords should be explained in Latin because it would provide a more learned appearance. However, the King, who himself was not an especially learned person, to put it nicely, wanted another solution. This was expressed in a letter which his Cabinet Secretary, Ove Høegh-Guldberg, wrote on November 28, 1776:


[The dictionary should be in the country’s own language and equally dear to the learned and unlearned. The natives should be those who have the greatest benefit from such a work. Foreigners should only turn over the leaves, our own should use it. The former may at most assess it in their pieces of information. On the contrary, our own should consult it as an adviser and get help to express themselves firmly and with certainty.]
Thus, the King, or his Cabinet Secretary, wanted a dictionary with definitions in Danish which could provide advice to both the learned and the unlearned Danish users when they needed help to express themselves correctly in their mother tongue. The Board of the Society of Sciences was not happy at all with the royal intermingling in what they considered their business, but as it was the King who was supposed to finance the project, it ended up in a sort of compromise where at least the definitions were written in Danish. The dictionary, with the stately title *Dansk Ordbog udgiven under Videnskabernes Selskabs Bestyrelse*, was eventually edited in eight volumes, the first of which was published in 1793 and the last one 112 years later, in 1905. Taking into account the epoch and conditions, and in spite of the conceptual differences between the first and the last volume, the dictionary was not bad at all, but it has suffered from a bad reputation in the academic world ever since, probably because of all the rumours created around what was considered an ill-timed and plebeian intervention from the Absolute Monarchy.

The problem of academic pride was especially acute in the Age of Enlightenment but it has never been completely eradicated and still survives in one form or another, at least within lexicography. This shows itself in many ways, e.g. in the contempt with which an encyclopedia like *Lademanns Leksikon* is treated because the explanations are not “scientific enough”, whereas the Danish National Encyclopedia, *Den Store Danske Encyklopædi*, is praised to the skies although many of its articles are only understandable to learned people and specialists. These reactions are indicative of two different approaches to lexicography. According to the function theory, which is defended at this Centre for Lexicography, dictionaries and other lexicographic works should necessarily be conceived and produced in order to satisfy the information needs of the target user group, in this case the whole Danish population’s needs for encyclopaedic knowledge. It goes without saying that the data included in lexicographic works should be correct, and in this light it may be possible to criticize some of the articles in *Lademanns Leksikon*, but this can in no way justify that the “scientifically correct data” included in *Den Store Danske Encyklopædi* is presented in such a way that it is at least partially inaccessible and incomprehensible for the less educated users who make up a considerable part of the intended user group.
Another way in which academic arrogance constitutes an obstacle to lexicography is related to the development of its theory. During the last hundred years or more, a small group of those working in lexicography has made an effective take-over bid on the discipline and proclaimed that lexicography should be considered as “applied linguistics”. Although it cannot be denied that the practical results of linguistics have contributed positively to thousands of dictionaries, with this claim they disregard thousands of other lexicographic works for the production of which no specific linguistic knowledge was or is needed except for the language skills that any speaker possesses in his mother tongue. However, the main problem is not the exhausting struggle for the “ownership” of the discipline, but the fact that this conflict is hampering the development of an independent theory of lexicography which can help us not only to understand what has been done until now in terms of all lexicographic works, and not only a part of them, but also to guide the conception and production of entirely new lexicographic tools that meet the real demands of the present-day and future society.

**Faster horses**

It is frequently claimed that relying on the Market is the best way to stimulate innovation. I don’t believe in such authoritarian statements. And as regards lexicography, it is a fact that the Market, with a very few exceptions, is hampering the innovative process and impeding the production of the needed lexicographic tools. Only the dictionaries expected to be profitable are produced. And as long as the ones that are published are sold with the necessary profit, the big publishing houses have very little motivation to invest in real innovative changes. This is for instance the experience we have in Denmark with Gyldendal’s famous “red dictionaries”.

The problem here is the very concept of user needs. In a market economy, this will generally correspond to what the customers demand and what they are prepared to buy and pay for. This is also the way in which many lexicographic user surveys treat the problem as I have shown in an article that will be published later this year (cf. Tarp 2008b). However, if you focus on what the users buy, what they look up in dictionaries and what they themselves say or think that they need, you will end up in a sort of vicious circle, because the users will only consult a diction-
ary when they themselves recognize that they have a problem, when they consider that this problem is lexicographically relevant and when they, due to their previous experience with dictionaries, expect to find a solution by means of a lexicographic consultation. In this way you will only discover the shadows of the real user needs, and this is definitely not the best starting point for innovative thinking. It is necessary to break the vicious circle and find a completely new methodology to approach the problem.

About a hundred years ago when Henry Ford introduced his Model T Ford, he was asked if he had consulted the people before creating this model. His laconic answer was:

“If we had asked people what they wanted, they would have said faster horses.”

Although Ford as a capitalist was a tough guy, he was at the same time an excellent and audacious inventor due to his ability to go beyond the usual boundaries and satisfy people’s needs in a completely new way.

If we return to lexicography, what is being introduced during these years is almost exclusively fast horses. This is especially clear when it comes to the present transition from printed to electronic dictionaries. In 99.9 per cent of the cases, these new electronic dictionaries and lexicographic works are more or less tailored to the same pattern as their printed relatives with the only exception that data access is generally much quicker due to the use of search machines and links. As such, they are faster horses. However, what is both needed and possible today is the introduction of lexicographic Model T Fords. Only a few models of this kind have been invented so far, e.g. *Wikipedia* and the Danish Idiomatic Dictionary *Ordbogen over Faste Vendinger*. There is little doubt that the future, and even the near future, will demand lexicographic Mercedes Benzes and Rolls Royces, and it is therefore urgent to leave the fast horses behind and generalize the introduction of the first generation of Model T Fords in order to pave the way for the more advanced models of the future.

Fortunately, not everything is left to the Market in Denmark. There is a long tradition, going back to the time of King Christian VII and the production of the first big National Dictionary, that the State gives money to some important lexicographic projects. This is extremely important because it allows the production of lexicographic works which
would probably not have been published without this support. However, it is a major problem that, with the exception of one big project, all the big money has been given to Copenhagen where only about 30 per cent of the Danish population lives. And it is an even bigger problem that the financed projects, as a rule, are not carried out according to an advanced lexicographic theory, but to the theories that the people involved have brought with them from their respective branches of knowledge, e.g. linguistics. This implies, among other things, that the projects drag on for too many years – such as *Jysk Ordbog* which has been in progress since 1932 – that the projects are much more expensive than necessary, and that the final product is not the lexicographic work needed by the Danish people as such, but by the elite and the few. This is the case with two recent projects of national importance such as *Den Danske Ordbog* and *Den Store Danske Encyklopaedi*, which are definitely not made in the spirit of Ove Høegh-Guldberg although they – compared with similar works in other countries – have been finished in a relatively short time. And finally, it is a problem that the money has not been invested in real innovative projects, but only in faster horses, probably inspired by the nearby Klampenborg Racecourse.

**The function theory in ten points**

What is needed is a theory which, on the one hand, can guide the conception and production of lexicographic tools adapted to the specific needs of each and every person in each and every relevant situation, and which, on the other hand, can guide the planning and management of lexicographic projects in such a way that they will be carried through successfully with the smallest possible staff and within the shortest possible span of time, i.e. using the smallest possible amount of money. The money saved can then be used to finance other lexicographic projects of national, cultural, social and economic importance.

In order to develop a theory with this potential, the vicious circle mentioned above has to be broken. And the way to do this is to leave, for a moment, the world of dictionaries, the lexicographic consultation process, the user surveys and the questionnaires asking people what they themselves think they need, and jump out into the real world and look for real and lexicographically relevant needs there. This is what the function theory, defended at this Centre for Lexicography, has done.
When we temporarily left the world of dictionaries, things began to happen. Things we expected, but also strange things that took us by surprise.

I will spare you the long version. However, it is always difficult to make an understandable synthesis of a complex theory, especially to an audience that is not familiar with the problems of lexicography. I will nevertheless try to do it in ten points, but you will have to excuse me if I miss something or if you feel that anything is treated too superficially. The following are the most important elements of the function theory:

1. Dictionaries and other lexicographic works are like food, clothing and cars, i.e. they are utility products or tools which, like all other tools of this kind, are, or should be, designed to satisfy certain human needs.

2. Lexicographic needs are not abstract needs, but are always related to specific types of users who find themselves in a specific type of social situation.

3. In order to determine the nature of the lexicographic needs in every case, it is necessary to make a profile of the intended user group and determine the relevant social user situation.

4. The most important elements are the user situations of which three main types have been studied up till now, i.e. the cognitive situation where the potential user needs or wants to add something to his existing knowledge, the communicative situation where the user needs to solve a problem related to an ongoing or future communication process, and the operative situation where the user needs advice and instructions in order to perform any kind of mental or manual action, e.g. to operate a machine. Recently, it has been discussed whether there is a fourth main user situation, the interpretive one, where the user needs to interpret signals and symbols in the surrounding world, but it is still too early to conclude anything in this respect.

5. Lexicographic needs, apart from being user and situation-oriented needs, are always needs for information.

6. Lexicographic works do not contain information which, like knowledge, can only exist in relation to the human brain, but lexicographically selected and prepared data from which the user may retrieve the information needed through a mental process that lexicography does not study.

7. The information retrieved can then be used for a huge variety of purposes in relation to the social user situations mentioned above.
Lexicographic works are not the only texts that contain data from which information can be extracted. What distinguishes lexicographic works from these other text types is that they are not designed to be read from one end to another, but to be consulted. Rapid and easy access to the relevant data is therefore of utmost importance to lexicography. On this basis, a completely new discipline, i.e. *information and data accessology*, is being developed with the potential to project itself far beyond the limits of traditional lexicography and play an important role in the so-called information society. It may be useful not only for lexicographic works in the narrow sense of the word, but for all types of texts that are at least partially used for consultation, e.g. text books, user guides, etc.

**Audacity to go beyond the usual boundaries**

A major problem in the present information era is not the absence of data from which the information needed can be retrieved, but the abundance of unstructured data which implies that a person who, for instance, uses one of the existing search machines in order to find some specific data on the internet runs the risk of suffering from information death, i.e. being suffocated by the overwhelming amount of data popping up. Two things are needed to solve this problem. The first is the selection or preparation of need-adapted data and this can only be done in a scientific way if both the user profile and the social situation in which the needs occur are taken into account. The second thing is to guarantee the quick and easy access to exactly this data. There is no doubt that modern lexicographic theory can contribute to the solution in both cases.

In this perspective, it can rightly be considered extravagant that there are lexicographers who sit down for years wasting a lot of energy and public money in order to produce a dictionary for the few. Instead we should be inspired by the Chinese Emperor who six centuries ago organized a team of lexicographers who in only five years created a monumental work that gave access to most of the relevant knowledge of that time. We cannot repeat the experience, but we can definitely get inspiration to create something similar with new methods and technology. The fundamentals of a theory that can guide a modern enterprise of this dimension are already there, although the theory as such has to be fur-
ther developed. In this respect, we can seek additional inspiration in the words of wisdom of one of the greatest geniuses of the 20th century, Albert Einstein who in 1938, together with Leopold Infeld, wrote:

“The formulation of a problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old questions from a new angle, requires creative imagination and marks real advance in science.”

Allow me, in order to finish this lecture, to quote another big personality, in this case Georges-Jacques Danton, one of the leaders of the French Revolution. He said something about revolutions in general, that has relevance not only to social revolutions as the one he was engaged in, but also to scientific revolutions and, more modestly, to any innovative thinking, namely that in order to make a revolution three things are needed:

“Audacity, more audacity, always audacity!”

I believe that anybody doing research should make these three words their motto. At least, they will be mine. As a researcher you need the ability and audacity to go beyond the usual boundaries and generate new ideas, even if they are not understood and welcomed in the beginning.

Thank you very much for allowing me to share these ideas with you.

Bibliography

Dictionaries


188


Jysk Ordbog. Aarhus: Peter Skautrup Centret for Jysk Diaлектforskning 1932–. www.jyskordbog.dk


Literature


Tarp, Sven 2008: Lexicography in the borderland between knowledge and non-knowledge. Tübingen: Niemeyer.

Tarp, Sven 2008b: Kan brugerundersøgelser overhovedet afdække brugernes leksikografiske behov? In LexicoNordica 15. (To appear)