The pedagogical dimension of the well-conceived specialised dictionary

Sven Tarp
Aarhus School of Business
st@asb.dk

Abstract

In the course of the past 30 years it has been acknowledged that research on pedagogical lexicography has to address three main issues: (i) development of new formats to cover both communicative and non-communicative situations; (ii) the layout of the information on the dictionary page, to facilitate the handling of information; (iii) what information is necessary for each type of learner at specific levels of proficiency. Although a big amount of specialised dictionaries have been produced during the last three decades, only some of them show the pedagogical value or dimension of general dictionaries. To overcome this gap the theory of lexicographic functions indicates a way forward to the design and production of the lexicographic quality products that the modern globalized world requires. This article illustrates some of the workings of this functional theory by commenting on some of the pedagogical value of *EDGT*, a two-way, bidirectional and multifunctional dictionary.

Key words: LSP lexicography; pedagogical dimension; functional theory; knowledge-related consultations

Resumen

*La dimensión pedagógica de un diccionario especializado bien concebido*

A lo largo de los 30 últimos años la investigación relacionada con la lexicografía pedagógica se ha situado en la vanguardia de todo lo relacionado con los diccionarios. Se ha puesto de manifiesto, además, que es necesario continuar el estudio de, al menos, tres cuestiones: (i) el desarrollo de formatos nuevos con los que cubrir las situaciones comunicativas y no comunicativas; (ii) la distribución de la información en el diccionario; (iii) conocer la cantidad de información que un diccionario debe tener de acuerdo al nivel de conocimiento de sus usuarios. Los diccionarios generales muestran un gran desarrollo, especialmente los diccionarios pedagógicos. La mayoría de los diccionarios especializados, sin embargo, parece que no han tenido en cuenta estos nuevos enfoques. En este artículo se afirma que la teoría de las funciones lexicográficas ofrece un soporte teórico adecuado para introducir una dimensión pedagógica en el diccionario especializado. Ilustra esta afirmación con un trabajo
1. Introduction

Although a lot of water has flowed under the bridges since the German lexicographer Harald Weinrich in 1976 straightforwardly declared that specialised dictionaries were "lexicographic suckers", the situation in this respect has not changed a lot. A big amount of specialised dictionaries have been produced during the last three decades but except for a few exceptions, especially those based on the modern theory of lexicographic functions (see e.g. Bergenholtz & Tarp, 2002, 2003, 2004), most dictionaries of this sort still suffer from a very low quality compared with their LGP relatives. For that reason it is difficult to speak about the pedagogical value or dimension of specialised dictionaries in general. However, analysing the few exceptions from the point of view of the functional theory of lexicography it is possible to extract elements that could merge in a future and not yet existing theory on the pedagogical dimension of specialised dictionaries.

2. The pedagogical dimension of specialised dictionaries

According to the functional theory, all theoretical and practical lexicographic work should be based on the dictionary functions which represent the assistance provided by a dictionary—by means of its lexicographic data— to a specific type of user in solving the specific type of problems related to a specific type of user situation. This vision of lexicography differs from the point of departure taken by the majority of user surveys which, even if they observe the users in a specific situation, only focus on the look-up strategies which may already be "contaminated" because of the well-known fact that the users only look up what they expect to find in a dictionary, i.e. what they know from experience that dictionaries may contain. By contrast, what is studied in accordance with the methodology of the functional theory are those problems popping up in a specific situation to which a solution can be given by means of lexicographic data. The theory therefore concentrates on the extra-lexicographic user situations of which two main types can be distinguished: the communication-related and the knowledge-related. The first one is related to...
communicative situations such as L1 and L2 text production, L1 and L2 text reception and L1-L2 and L2-L1 translation. The second one is related to non-communicative situations where the users for one or the other reason want or need knowledge of a certain type, e.g. linguistic, general-encyclopedic or specialised knowledge. In the first case, the users –as a starting point– are only interested in solving problems that may arise during the communication process in order to make it go as smoothly as possible, whereas they, in the second case, only are interested in adding to their existing knowledge. It is in this last case that the specialised dictionary may display a pedagogical function providing LSP or specialised information to the user.

However, although on the one hand the two main types of user situations are quite different from each other, on the other hand, a complex dialectical relation exists between them. In the case of the knowledge-related user situation where the user wants to retrieve information from the dictionary that can be rooted as knowledge, a new type of passive communication (time-delayed text reception) is established between the user and the original dictionary author, which may imply communicative problems that have to be solved by means of the same or another dictionary. And if, in the case of the communication-related user situation, the information extracted from the lexicographic data contained in the dictionary takes root in the user's brain or memory, then it is transformed into knowledge that can be used in future communication, and the same thing happens if the consultation related to a communicative problem arises the user's interest in the topic represented by the lexicographic data and turns into a knowledge-related consultation and exploration of the data, thus transforming the original purpose of the consultation and replacing it with another one. It is evident that the dialectical relation between these two fundamental results of a dictionary consultation should not be confused with the relation between encyclopedic and linguistic knowledge as it is the case in Wiegand (1988, 1994) for which he has been criticized by Rossenbeck (1994) and Bergenholtz (1998). Knowledge about the world cannot be stored without language, and linguistic skills are not sufficient in order to engage into a meaningful communication as this also presupposes general or specialised encyclopedic knowledge and both may be achieved fully or partially through the communication-related consultation whereas each of the two main types of knowledge-related consultations, in principle, only partially furnishes elements to be used in a meaningful communication. But even between these two types of knowledge-related consultations no Chinese Wall is to be
found. Encyclopedic knowledge cannot be imparted without language and this process includes the transmission and possible assimilation of terminology and part of the corresponding grammar. And if the imparting of linguistic knowledge, apart from formal grammatical knowledge about morphology, syntax and word formation (e.g. the use of affixes), also includes terminology, then a transmission of elements of encyclopedic knowledge will inevitably take place between the author and the user with the dictionary as the medium.


The pedagogical value of a dictionary is thus determined by its appropriateness to provide either linguistic or encyclopedic knowledge, or both, to the users during the process of linguistic or encyclopedic learning. Consequently, if the pedagogical value of a specialised dictionary is determined by its appropriateness to provide LSP and/or specialised encyclopedic knowledge, then every dictionary which contains data representing true knowledge may, up to a certain point, display a pedagogical dimension, although only the dictionaries designed for this specific purpose would be of any real benefit to the users interested in assimilating specialised linguistic and encyclopedic knowledge. In order to explore the pedagogical possibilities of a well-conceived specialised dictionary in this respect, a concrete pair of dictionaries where the authors have had the rare editorial freedom to implement all their ideas, will be discussed in the following. These two volumes, which actually constitute the two language directions of one and the same dictionary, are the Encyclopedic Dictionary of Gene Technology English-Spanish and the Diccionario Enciclopédico de Ingeniería Genética Español-Inglés, in the following abbreviated EDGT. This dictionary is a two-way, bidirectional and multifunctional dictionary. It embraces, as already said, both language directions, it is designed to users with both English and Spanish as their mother tongue, and it is conceived to provide assistance to various types of users in various user situations as can be read in the preface:

It has been our aim to design a dictionary which can serve as a tool for those who will be or are in contact with gene technology in some way and who wish to acquire
a certain or an additional knowledge for the field and its language through the
dictionary. More specifically, this group could be public sector or private sector
employees who require an aid in their work in connection with changed methods of
medical treatment, patent applications, food inspection, environmental issues etc.
Thus, the users of this dictionary consist of a wide range of persons whose
background may be in medicine, technology, pharmacy, biology and/or journalism.
In addition, translators translating texts in which terms from the field of gene
technology occur may benefit from using the dictionary as well. [...] Our dictionary
provides both purely factual information and specialised linguistic information on
gene technology. Thus the reference work is intended to be an aid to understanding
purely encyclopedic relationships and an aid in the communication of topics of gene
technology. (EDGT, 1998: 5)

In a special component part called "Information About the Structure and
Background of the Dictionary", the eight main user situations to which the
dictionary is designed are specified:

1. reception of Spanish texts
2. production of Spanish texts
3. reception of English texts (modificado)
4. production of English texts (modificado)
5. translation from Spanish into English
6. translation from English into Spanish
7. information on subjects within molecular biology
8. indirect text-related information about Spanish and/or English molecular
language use (EDGT, 1998: 7)

As the intended user group is made up by both native speakers of English and Spanish
that might be experts, semi-experts and so-called "interested laymen" within the sphere
of gene technology, the number of possible functions of the dictionary raises
considerably although not all types of users may need the dictionary in each and every
of the eight situations mentioned above. The interesting user situations from the point
of view of this article are the knowledge-related situations, i.e. number 7 where the users
may consult the dictionary in order to get specialised encyclopedic knowledge (on gene
technology) and number 8 where they may consult it in order to get the corresponding
LSP knowledge that is not related to a specific communications situation.
However, one thing is the intended functions of the dictionary; another thing is the authors’ success in terms of compiling a lexicographic product that lives up to its declared purpose. Two things have to be taken into account in this regard, first the types, amount, quality and relevance of the data fixed in the dictionary items and, secondly, the principles according to which these data are structured, connected and presented in order to facilitate their accessibility. The data contained in a dictionary can be of two different types, i.e. the function-related data that provide solutions to the problems that originally gave raise to consultation, and the use-related data that provide information on the dictionary and assistance in using it. Both types of data are important in order to generate a quality product and can be split up and structured in separate component parts. *EDGT* consists of eleven such components parts:

1. Preface (in English)  
2. Preface (in Spanish)  
3. Information About the Structure and Background of the Dictionary (in English)  
4. Information About the Structure and Background of the Dictionary (in Spanish)  
5. Literature consulted in the preparation of the dictionary (repeated in both volumes)  
6. Guide to the Use of the Dictionary (in English)  
8. Introduction to Molecular Biology (in English)  
9. Introduction to Molecular Biology (in Spanish)  
10. English-Spanish Word List A to Z  
11. Spanish-English Word List A to Z

The first seven are use-related component parts whose purpose it to explain the origin and functions of the dictionary and provide guidance to the users, whereas the last four are function-related component parts. In the following only the last ones will be discussed.

The only component part that is compulsory and included in the very definition of a dictionary is the word list. Any other component part is facultative, although most dictionaries, apart from the central word list, also contain other component parts, especially use-related component parts such as user guides and prefaces. However, depending on the intended user group and user situations, a number of dictionaries also place part of the function-related data in special component parts that may be integrated (by a system of references) or non-integrated into the central word list.
This distribution and structuring of the lexicographic data within a specific component part and between various component parts are determined by the data distribution structure.

As mentioned above, EDGT contains a separate systematic section called "Introduction to Molecular Biology" and provided in both an English and a Spanish version which are integrated into the two central English-Spanish and Spanish-English word lists by means of a complex system of cross-references. Such integrated function-related component parts are especially appropriate for dictionaries with knowledge-related functions and can be used even by persons with little knowledge of gene technology. According to the authors, only "a basic knowledge of science and an interest in the field's scientific processes and results are […] presupposed in order to benefit from the dictionary" (EDGT, 1998: 7). Through this systematic introduction to molecular biology (or gene technology), it is possible to study and assimilate (learn) the basic principles and terminology of this specific subject field. The section consists of a total of 43 pages and includes a table of contents referring to 30 paragraphs organised in four chapters:

**From Chemistry to Biology**

§ 1. Molecular Biology 20

§ 2. The Chemistry of Cells and the Origin of Life 20

– The Origin of Life 22

§ 3. Prokaryote 23

§ 4. Eukaryotes 23

§ 5. Mitochondria and Chloroplasts 24

§ 6. Membranes and Cytoskeleton 25

§ 7. Chromosomes 26

§ 8. Some Properties of Bacteria 27

§ 9. Some Properties of Eukaryotes 28

– Cell Divisions and Life Cycles in Eukaryotes 30

– Yeast, a Unicellular Eukaryote 31

– The Immune System 31

– Respiration and Photosynthesis 31
The Small Biomolecules
§ 10. Water and Organic Compounds 32
§ 11. Carbohydrates 33
§ 12. Amino acids 34
§ 13. Nucleotides 35
§ 14. Lipids 35

The Large Biomolecules
§ 15. Macromolecules 36
§ 16. Proteins 37
§ 17. Polysaccharides 38
§ 18. Nucleic Acids 39
  – Viruses 41
  – The Central Dogma of Molecular Biology 42

Recombinant DNA and Cloning
§ 19. DNA Technology 43
§ 20. Restriction Enzymes 43
§ 21. Vectors 44
§ 22. DNA Libraries 47
§ 23. Probes and Screening of Libraries 47
§ 24. Chromosome Walking 49
§ 25. Characterization of Genes and Their Products 49
§ 26. Expression Vectors 51
§ 27. Reverse Genetics and Protein Engineering 51
§ 28. In vitro Cloning of DNA 52
§ 29. Genome mapping and DNA Fingerprinting 53
§ 30. Gene Technology and Society 55
  – Preamble 55
  – Code of Conduct 57
  – The Code 58

(EDGT, 1998: 19)

Text example 1: List of content of the "Introduction to Molecular Biology" in EDGT
This table of content provides in itself a general and systematic overview of the whole subject field of molecular biology. Consequently, the text in the introduction is organised in a systematic way according to the system of the subject field. It is illustrated with a total of 23 figures and the most important terms are written in bold and serve, in addition, as indirect cross-references to the central word lists where a more detailed explication can be found:

§ 1 Molecular Biology
Molecular biology is the study of molecular aspects of biological phenomena. In molecular biology the structure and function of cells, their components and their interactions are investigated by physical, physicochemical and biochemical techniques, and the results are interpreted in terms of biomolecules, especially macromolecules, and their interactions. Molecular biology encompasses and integrates many molecular aspects of genetics, physiology, microbiology, biochemistry and cell biology.

§ 2 The Chemistry of Cells and the Origin of Life
The cell is the smallest structural and functional unit of living organisms. The smallest organisms consist of single cells, whereas larger organisms are multicellular. The higher multicellular organisms contain many different types of cells with specialized functions. The various cell types cooperate in a tightly coordinated way, although each cell retains a certain independence.

The plasma membrane defines the periphery of the cell. It is composed of a large number of lipid and protein molecules and is very flexible, allowing changes in the shape and size of the cell. All cells, with a few exceptions, contain a genome, i.e. a complete set of genes, consisting of DNA. A few highly differentiated and specialized cell types lose part or all of the genome. As an example red blood cells have no DNA and no cell nucleus.

There are two main types of cell, the prokaryotic cell and the eukaryotic cell. Bacteria and cyanobacteria, or blue-green algae, are prokaryotes, and animals and plants consist of eukaryotic cells. In eukaryotic cells the genome is located in a cell nucleus and is therefore separated from the cytoplasm by the nuclear envelope. In prokaryotes the DNA is not surrounded by a nuclear envelope membrane, but is located in a more or less distinct region of the cytoplasm called the nucleoid. The diameters of bacteria are typically 1-2 µm (1 µm = 10⁻⁶ m), and most eukaryotic cells have diameters of 10-100 µm.
The internal volume bounded by the plasma membrane, the cytoplasm, is composed of a complex and gel-like aqueous solution, the cytosol, and a variety of suspended particles. Dissolved in the cytosol are many enzymes and other proteins and the RNA molecules that encode them, the monomeric subunits, amino acids and nucleotides, from which these macromolecules are assembled, many small organic molecules that are intermediates in biosynthetic or degradative pathways, and a variety of inorganic ions. **Ribosomes** are complexes of more than 50 proteins and RNA molecules with a diameter of about 0.02 µm. They are the sites where protein synthesis, **translation**, occurs. Plant cells are surrounded by a cell wall and contain a range of distinct membrane-bounded structural and functional units or organelles, e.g. mitochondria, vacuoles, chloroplasts, and endoplasmic reticulum. Animal cells lack cell walls and chloroplasts but apart from that the basic structures and functions of animal and plant cells are generally identical. Prokaryotes do not contain membrane-bounded organelles. *(EDGT, 1998: 20)*

**Text example 2: Extract from the "Introduction to Molecular Biology" in EDGT**

There are three ways to make use of the introductory section. The users can read all of it from the beginning to the end and thereby get a systematic introduction to molecular biology in general, and the users can also choose only to read a chapter or just one or a few paragraphs which are most relevant for them. If they opt for this last procedure, they can access the relevant paragraphs both through the above table of content and by means of cross-references from the central word lists. And if the users, during the reading of the systematic section, need additional information, then they can access the relevant articles in the central word lists through the indirect cross-references provided by the subject-field terms written in bold. The articles (entries) have the following internal structure:

**translation traducción f**

The biosynthesis of protein is termed translation. The basic principle of the process is that the amino acid sequence of the protein is determined by the genetic code i.e. the nucleotide sequence of the gene, which temporarily manifested as mRNA, will be expressed in the protein’s primary structure, the amino acid sequence.

Translation can be divided into two simultaneous processes.

1. The transmission of the genetic code from mRNA to the amino acid sequence of the protein. The basic principle is that the triplets of the mRNA determine
the aminoacyl-tRNA that can be accepted because of the specific binding of the triplet to the tRNA anticodon.

2. The biochemical reactions which establish the chemical bonds between the amino acids, the peptide bonds.

The ribosome is an important structure in translation. Its two subunits form a chemical complex together with the mRNA at the initiation site, the ribosome binding site. The aminoacyl-tRNA molecules bind to two sites in the ribosome mRNA complex, first to the aminoacyl-tRNA, the A-site, and after that to the peptidyl-tRNA site, the P-site. This results in the formation of the peptide bond. The unloaded tRNA, i.e. tRNA which has delivered its amino acid, leaves the ribosome complex and simultaneously this complex moves one step, jumping a triplet length on the mRNA. Now a new aminoacyl site is opened for the anticodon of the next aminoacyl-tRNA and the second amino acid is coupled to the first one. Chain elongation continues. The movement of the ribosomes makes space for a new initiation complex behind the first one. In other words several ribosomes can move along the mRNA, producing several protein molecules at the same time. Reaching the termination signal i.e. the stop code the ribosome releases the mRNA and dissociates into its two separate subunits, which can assemble on another mRNA molecule to begin a new round of protein synthesis.

The production of the primary protein has been completed and posttranslational modifications of the primary molecule often take place.

- The genes encoded in the double helix control cellular activity through two processes: transcription, in which DNA molecules give rise to RNA molecules, followed by translation, in which RNA directs the synthesis of proteins.

\[\text{arrest protein} \sim \text{detener la} \sim \text{de las proteínas}; \text{carry out} \sim \text{llevar a cabo la} \sim; \text{halt} \sim \text{by binding to mRNA} \sim \text{detener la} \sim \text{mediante enlace con el ARNm}; \text{promote} \sim \text{promover la} \sim\]

\[\text{cell-free translation; transcription; genetic code; ribosome; ribosome binding site; \S\ 16}\]

Text example 3: Synopsis article from \textit{EDGT}

This article is a so-called synopsis article that gives an overview of the specialised knowledge related to a number of articles and the corresponding terms. In this article, the user can get more profound and special information about the two meanings of the polysemic term \textit{translation} and, furthermore, cross-references (after the symbol \(\Rightarrow\)) to other relevant articles as well as to a paragraph (\S\ 16) in the
systematic section that furnishes more general information about the term and its place in the systematic classification of the subject field. In addition, the article provide information about part of speech, collocations (after the triangle ▲) and even an authentic text example from the corpus (after the bullet •). The word list contains a number of such synopsis articles, but it also contains two other types of articles, individual articles which only includes encyclopedic data directly addressed to the lemma, and reference articles which only gives a reference to another article where the relevant encyclopedic data are to be found:

**translation product** producto m de la traducción

The primary translation products are almost always precursors to functional proteins. They are altered by posttranslational modifications.

→ translation

**translation termination** terminación f de la traducción

→ translation; stop codon

Text example 4: Individual article and reference article from *EDGT*

### 4. Benefits of *EDGT* for Different Types of Users

As can be seen, there is a certain overlapping between the data placed in the different articles in the word list and between these data and those of the systematic section. In fact, the users can find data at six different interconnected levels:

1. reference articles which only provided references to other articles
2. individual articles with data addressed to the lemma (term)
3. synopsis articles which gives an overview of data related to various terms
4. paragraphs in the introduction to molecular biology
5. chapters in the introduction to molecular biology
6. the introduction to molecular biology as a whole

This type of data distribution structure which is made visible to the users by means of a complex system of cross-references is particularly suitable when the specialised dictionary has to display a pedagogical function providing special information that can be transformed into knowledge about the subject field and its terminology during the
learning process. It offers a great versatility and accessibility that permit different types of
user with problems related to different types of user situations to start their dictionary
consultation at different levels that correspond to their specific needs. And the possibilities
put at the disposition of the users are even bigger because of the fact that all the six levels
mentioned can be accessed both from Spanish and English thus providing a unique
possibility to study and compare also the Spanish and English molecular biological LSP.

In the following, the benefit which four different types of users can draw from the
dictionary in different user situations in terms of LSP and subject-field information will be
discussed. The first type is a journalist which is among the intended users mentioned in the
dictionary preface. If the journalist for example wants to write an article on any topic
related to a company using gene technology in the production, then the consultation of the
systematic introduction, or part of it, may be of great benefit in order to get a general idea
of the technology applied and of the correct terminology, i.e. to retrieve both encyclopedic
and linguistic background information that may be necessary in order to prepare and write
a well-documented and serious article on the company. Besides, the journalist may also use
the dictionary in a communication-related situation, e.g. when he or she is reading another
text about gene technology and comes up with a number of hitherto unknown terms which
can be explained and understood through a look-up in one of the two word lists. However,
this kind of dictionary consultation is not primarily made in order to generate knowledge,
but to solve concrete problems popping up during the text-reception situation. Although
the information retrieved from the dictionary may be rooted in the journalist's memory and
thus transformed into lasting knowledge, this is not among the predicted pedagogical
functions of the dictionary which are mainly related to the knowledge-seeking consultation.

Another type of user who is also mentioned in the preface is a pharmacist, e.g. doing
research within his or her own subject-field. The pharmacist can be described as a semi-
expert within the field of gene technology because there is a number of overlapping
disciplines between pharmaceutics and molecular biology. However, the pharmacist is not
a complete expert of gene technology and might therefore benefit from the dictionary in
terms of adding to his or her existing knowledge. This consultation could start in the
systematic introduction or even in the word lists depending on the concrete interest and
needs of the pharmacist.

A third type of user may be a translator preparing for a translation job related to gene
technology and not being a specialist of the field. As a qualified translation of specialised
texts dealing with complex topics as gene technology cannot be made without at least a certain understanding of what the text is all about, the consultation of the systematic introduction, or the relevant parts of it, may be of great assistance in the preparation for the translation. And the translator may also be able to assimilate the fundamental terminology in both English and Spanish by reading the introduction in both languages.

A fourth type of user could be the very expert in gene technology who is preparing a paper or conference in a language which is not the one that he or she normally uses to communicate in terms of the subject field. This kind of user may be a native English speaker preparing a paper or conference in Spanish, a native Spanish speaker preparing a text in English and even a native Spanish speaker preparing a text in Spanish as he or she might not know the corresponding LSP because of the fact that gene technology in many or most cases is studied and communicated in English which is the lingua franca in terms of this subject field. Thus, the reading of the introduction in the "other" language and making use of the relevant cross-reference may furnish the necessary vocabulary in order to write the paper or prepare the conference.

5. Conclusion

For all these four types of users—as well as for others of the intended users—the dictionary displays an important pedagogical function that allows it to meet their specific type needs in a specific type of knowledge-related user situation. This result of the authors' work has only been achieved by applying the theory of dictionary functions to the practical specialised lexicography. However, although there are few such dictionaries, EDGT is not the only dictionary designed according to these theoretical principles. Among other important dictionaries are:

- Gubba (1995) which contains a parallel introduction to the German and Danish legal systems written by Jan Engberg and especially dedicated to translators that have to be acquainted with these systems in order to avoid pitfalls and produce a correct translation in terms of its legal content.
- Henriksen (2005) which provides a systematic introduction focussing on the differences between the Danish and Spanish legal systems and with the same purpose as the above mentioned dictionary.
- Andersen et al. (2005) which includes a separate contrastive grammar of English and Danish for business, written by Birger Andersen, that permits
the users to study and assimilate the peculiarities and differences between the English and Danish LSP independently of any specific communication-related user situation.

If specialised dictionaries should leave behind the period of "lexicographic suckers" and not only be fully adopted to the communication-related user problems, but also display a pedagogical function in terms of knowledge-related user situations, then there is little doubt that the principles and practice discussed in this article indicate a way forward to the design and production of the lexicographic quality products that the modern globalized world requires.

(Paper received June 2005)

REFERENCES


Sven Tarp is an authorised translator and interpreter with a PhD in specialised lexicography. He is an associate professor of Spanish at the Aarhus School of Business, Denmark, and the editor, author and co-author of a large number of dictionaries and theoretical articles and books on lexicography.