Dictionaries in the Internet Era:
Innovation or Business as Usual?
(*Enrique Alcaraz Memorial Lecture 2014*)

Sven Tarp
Centre for Lexicography
Aarhus University, Denmark
st@asb.dk

ABSTRACT

This article is the written version of the Enrique Alcaraz Annual Memorial Lecture, given by the author at the University of Alicante on March 26, 2014. After a brief overview of the past and present of lexicography, it presents some of the challenges and paradoxes facing the discipline in the current transition to the digital media. Through examples from lexicographical practice it provides a vision of dictionaries as information tools and presents the core elements of a general theory covering all types of lexicographical work. Upon this basis, the article discusses principles, methods and techniques which can be applied to improve the quality of present and future online dictionaries. Finally, it presents the main concept of an online business dictionary under construction where some of these principles, methods and techniques are used in order to adapt the articles visualised on the screen to the foreseen Spanish and English users' needs in five different situations.

Keywords: e-Lexicography, e-dictionaries, function theory, information overload, lexicographical challenges

1. Introduction

Thanks to the University of Alicante, the Department of English Philology and Professor José Mateo Martínez for the invitation to participate and speak at this Annual Memorial Conference, organised in honour of the late Professor Enrique Alcaraz Varó.

I only met Enrique Alcaraz once, during the Winter University organised in Soria in November 2005. Nevertheless, I remember him perfectly, as a nice and pleasant man with a broad culture who was able to entertain his colleagues for hours due to his big encyclopaedic knowledge. A few years later, I was asked to contribute with two articles to a book about Specialised Dictionaries for Learners, dedicated to his memory and edited by Pedro A. Fuertes-Olivera (Tarp, 2010; Bergenholtz & Tarp, 2010). And now, much to my surprise, and to my great pleasure, I have been invited to address this event, also held in the memory of this respected former Professor of the University of Alicante.

Enrique Alcaraz was a distinguished scholar with a broad academic production within various fields, among them lexicography where he was the author of a number of fine specialised dictionaries, known and respected not only in Spain but also abroad. Hence, it seems quite logical to choose lexicography as the topic of my lecture today.

Lexicography is a several thousand years old social and cultural practice. It has, at least, been practised since the third millennium before our era when Sumerian scribes wrote cuneiform word lists by pressing reed tools into clay tablets and thus producing the first known dictionaries. Since then, hundreds of thousands of different dictionaries have been compiled; nobody will ever know their exact number. Dictionaries with many different names; beloved child has many names. Dictionaries carved in clay, hand-written on papyrus, printed on paper with various techniques, or made available on one of the many digital platforms existing today, notably the Internet. Dictionaries in almost all languages, as well as many dialects. Progressive dictionaries promoting the advance of humankind and, sorry to say it, also reactionary ones trying to stop the wheel of history. Small dictionaries with only a few dozen words; and huge dictionaries with hundreds or even thousands of volumes, the biggest one ever produced with no less than 11,095 volumes (the *Yongle Dadian* from 1408). In short, there are dictionaries for any taste.

One of the most fascinating things about dictionaries is that they, during the millennia, have covered almost any area of human activity. As such, dictionaries constitute a privileged mirror of social and cultural development during the past four thousand years, not only in terms of the development of languages, but also of handy-craft, economic life, culture, education, natural and social sciences, humanities, sport, and even such exotic phenomena as entertainment, pastime, holiday, etc. Just as dogs have been man's best friend for more than thirty thousand years, dictionaries have been man's faithful companion for about four thousand years. If you can see the wood for threes, it is a privilege to work within this field.

2. Challenges

In a few years lexicography has been whirled into what some have called the "information society" and others the "mis-information society" (cf. Robins & Webster, 1983). As a result, this old and fascinating discipline is now passing through a complex process which in a certain sense could be interpreted as a sort of crisis, or even an identity crisis. There is no risqué that the present problems should threaten its existence as such, but they may somehow challenge its role in society. The situation is as follows: on the one hand, dictionaries have reached a bigger audience than ever before thanks to the new media; and on the other, they experience a relative loss of users who prefer consulting other kinds of information tool in order to get the answers they need. There are various reasons for this dual development which will probably accelerate in the nearby future. In the last analysis it is detonated by the new computer and information technologies and techniques introduced into lexicography during the last decades, starting very modestly in the mid-sixties of the last century. This phenomenon has led to a revolution – or at least the need for a revolution – in almost all aspects of practical lexicography. This is not only true in regard to the presentation of the final product to be consulted by the users, a product which is increasingly placed on an electronic platform; it is also the case in most of the operations related to the compilation of this product as well as to the research into its usage.

When a millenarian culture practice like lexicography takes the gigantic step from one platform to another, i.e. from the printed to the digital media, then one would expect this transformation to be much more than a mere change of form. One would indeed expect a revolution also in terms of quality which in lexicography can be translated into a better and more individualised satisfaction of user needs. However, various facts seem to indicate that the "old man" is poorly dressed to confront the current climate change. Lexicography has showed that it is not immune to the new plagues created by the information society. Old habits already problematic in the printed world have been transferred to the virtual universe. A mayor problem is that the extremely rich experiences from this millenarian social and cultural practice have not been summed up with the necessary scientific rigor. In fact, when looking back with self-critical eyes and discovering its real core and essence as the world's probably first information discipline, lexicography is in an advantageous position to provide answers to many problems observed in modern (mis)information society. Let us look at some of the present problems.

3. Paradoxes in modern lexicography

One of the paradoxes in modern lexicography is that the new technologies introduced do not only bear the promise of solutions to many of the problems observed in existing dictionaries, but also themselves add to these problems. In a personal comment on the future of learner's dictionaries, Michael Rundell, the Editor-in-Chief of the Macmillan

family of English learners' dictionaries, writes the following about the eighth edition of the *Oxford Advanced Learner's Dictionary* (OALD):

> The current business model is not sustainable. Already, the books have become bloated almost to the point of impracticality, as publishers add more features with each new edition. The latest contender, OALD8, comes in at just under 2,000 pages [...], over twice the size of OALD3. (Rundell, 2010: 170)

One of the main reasons for this complex and contradictory development of printed learner's dictionaries – and, up to a certain point, of printed dictionaries in general – is the introduction and generalized usage of big electronic corpora which allow the lexicographers, in a much easier and quicker way than ever before, to retrieve a lot of relevant lexicographical data which may be helpful to the intended users in one situation or another. As could be expected, the dark side of the picture is that these data have bloated the printed dictionaries "almost to the point of impracticality". Although a passable way out of this blind alley may seem to be the conversion to electronic dictionaries connected to databases with a huge storage capacity, this step is apparently not as simple as it looks. For instance, in the same book as Rundell, Wendalyn Nichols, the Commissioning Editor for Cambridge Dictionaries, writes that "the true advantages of Web-based dictionaries" are "freedom from the space constraints of the printed book and on-demand updatability" (Nichols, 2010: 40).

There is little doubt that on-demand updatability is an advantage of Web-based dictionaries in comparison to their printed relatives. At the Centre for Lexicography in Aarhus, for instance, a number of both general and specialised online dictionaries, monolingual as well as bilingual, are updated almost daily, based on either comments from the users or study of log files showing that users have tried in vain to look after words which are relevant for the dictionary in question though still not treated in it. However, when it comes to the "freedom from space constraints" which Nichols also mentions, then it is worth noting that a printed dictionary may become bloated in two dimensions, vertically and horizontally, requiring very different solutions:

- *Vertical bloating*, i.e. the growing number of lemmata.
- *Horizontal bloating*, i.e. the ever increasing amount of lexicographical data attached to each lemma.

Vertical bloating is a problem related to the printed universe as it can easily be surmounted in an electronic dictionary connected to a database with sufficient storage capacity. Horizontal bloating, however, constitutes a much more serious and complex problem whose solution requires a completely different approach. It is related to the phenomenon known in information science as *information overload*, the famous concept introduced by Miller (1956) and popularised by Toffler (1970). Today, information overload has become an international plague, especially on the Internet. It is sufficient to recall what happens when one uses Google's search engine in order to obtain specific information for whatever purpose and then is referred to thousands of

pages, of which the first ones are often annoying, irrelevant publicity paid by big business. Sometimes you are lucky, but frequently – especially if you are in a hurry – you just get lost in the huge amount of irrelevant data appearing on your screen. In lexicography, this problem expresses itself in the presentation of much more data than required by the user in a given consultation. Such an overload of lexicographical or other kinds of data may delay or even hinder the access to the relevant data as well as the retrieval, from these data, of the information needed. In the following we will look at four examples of how this phenomenon expresses itself in the "real lexicographical world", in all cases acting as the devil's advocate.

The first example is related to Oxford University Press, the prestigious British publisher of a big number of dictionaries of a great variety of types. Many of these dictionaries have been placed totally or partially on the Internet through the web portal *Oxford Reference*. However, from this portal it is very difficult, if not impossible, to get access to the individual dictionaries, meaning that users with a lexicographical need in most cases have to make use of the search engine of the portal. If, for instance, the word *table* is written in the search field, then it will give a result with reference to no less than 12,511 entries in different dictionaries (see Example 1).

Example 1: Search results in Oxford Reference

This is a typical example of data overload in the access phase. The user will either have to spend a lot of time, which could be used for much more productive purposes, before obtaining the required information or will simply get lost and interrupt the search process. And if the user finally gets through to the corresponding article in, for instance, the world famous *Oxford English Dictionary* (OED), he or she will find that the article *table* contains about 35,000 words which represent more or less 90 pages in a MS Word document. Example 2 shows the information which OED provides "about" itself.
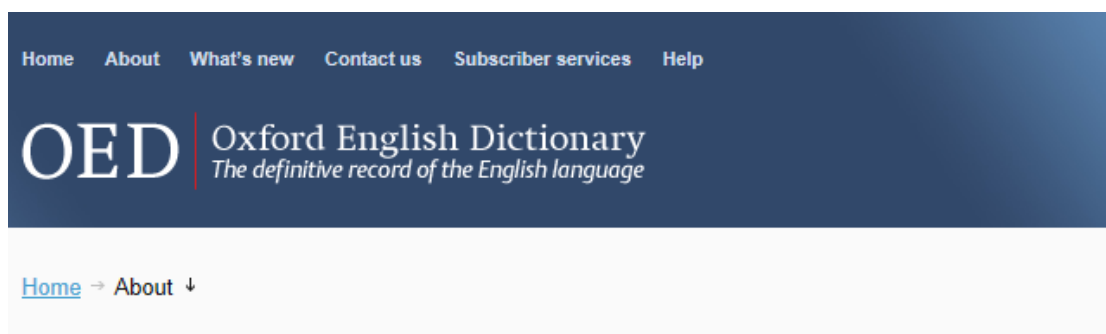
Example 2: About Oxford English Dictionary

This widely accepted authority on the English language is a marvellous source of information if one wants to study the English language in details and know something about its history and development. But it is definitely not the most adequate tool to assist a user in need for a quick answer when having, for instance, text-production or reception problems, although many people do use it with this purpose due to its well-earned prestige in other respects.

The two examples from Oxford are, of course, extremes used to illustrate the problem of information overload in various phases of the consultation process. In most cases, this problem expresses itself in more "modest" ways, and not necessarily through

data provided directly by the lexicographers. In many free-access dictionaries, publicity is the only way to obtain profits to the publisher and finance the maintenance of the web page and the underlying database as well as the lexicographers' salaries, if any. The *BusinessDictionary.com*, consulted by many business people, illustrates this problem (see Example 3 where the dictionary has been consulted from Denmark).
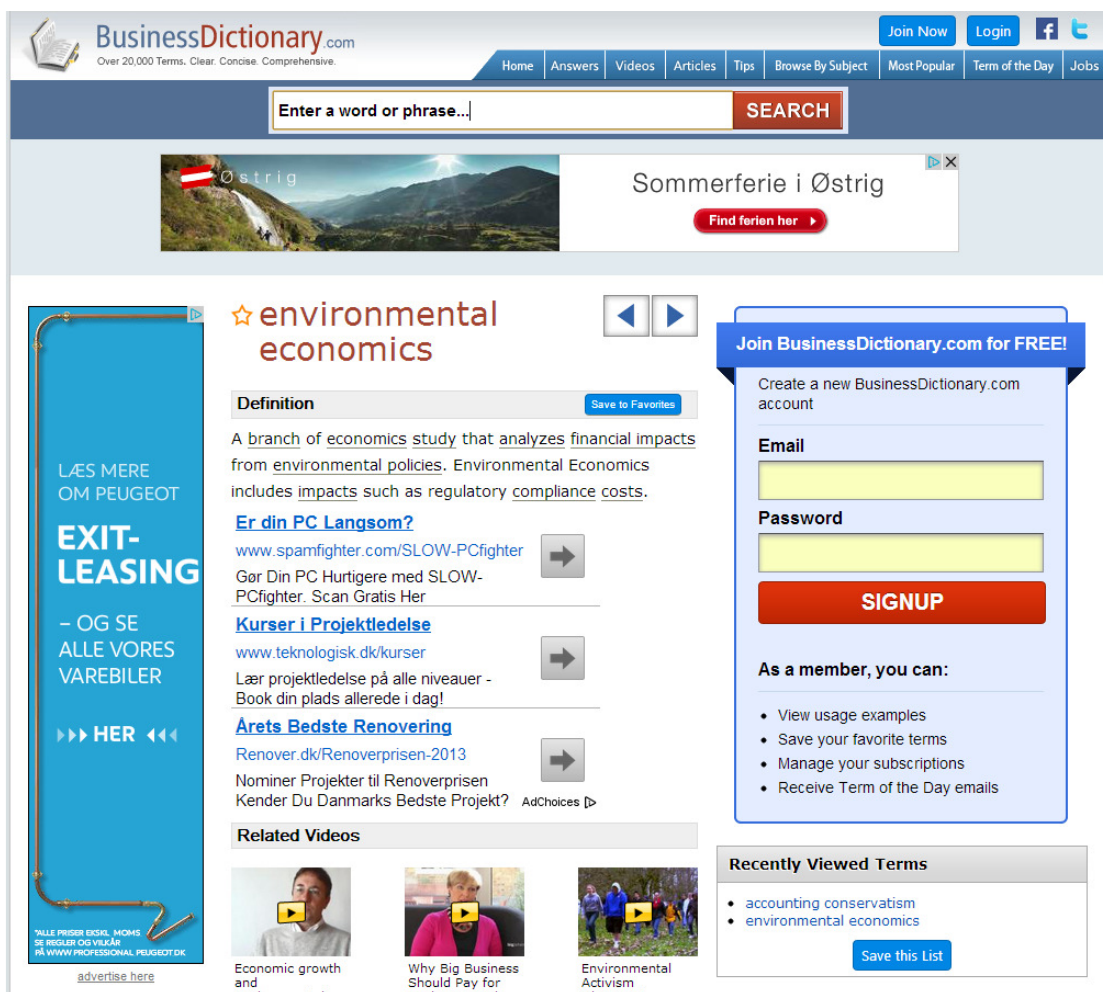
Example 3: Article from Business Dictionary

In this Internet dictionary, the lexicographical data required by the users in order to retrieve the needed information on the term *environmental economics* is almost hidden in the middle of a big accumulation of other, in this connection, completely irrelevant data. The publicity companies know what they are doing. From their research into user habits they know that the users consulting this dictionary will, although in most cases briefly, have a glance at their publicity with the desired effect. However, from a lexicographical perspective the effect is entirely negative as it complicates the consultation process in terms of data access and information retrieval. The only logical solution to this problem would be that these socially highly relevant tools were socialised and supported by public funding for the benefit of the public. But such a

solution, although lexicographically relevant, is basically a political decision to be taken outside the sphere of lexicography in the narrow sense of the word.

A final example showing the negative consequences of information overload, among others, is that of a professional translator of specialised texts who needs to make, say, 50 dictionary consultations in a normal workday. Such a number of consultations are completely normal, especially when it is a question of translating specialised texts containing subject-field knowledge unknown to the translator. However, if each consultation requires an average of 5 minutes due to access problems, information overload and other difficulties, this means 250 minutes spent in dictionary consultation, which for most people make up more than half a workday. This constitutes some rather expensive production costs for the translators and, as such, a terrible waste of time, cf. Nielsen (2008). If the average consultation time could be reduced considerably, this would immediately lead to higher productivity and less secondary production costs for the translator. In this respect, the time factor – expressed in quick and easy data access and information retrieval – is in itself an important criterion of lexicographical quality.

If we return to the "freedom from the space constraints" emphasised by Nichols (2009), Lew (2014) points to the ambiguousness of the notion of *dictionary space* and proposes a preliminary differentiation between *storage space* and *presentation space*:

> On careful inspection, it appears that the notion of *dictionary space* is not specific enough as a technical term, because it is ambiguous. The suggestion that dictionary space is unrestricted is actually largely correct, but only when space is understood as the capacity to hold the total content of the dictionary – this sense of *dictionary space* could provisionally be called *storage space*. There is at least one more important sense of dictionary space which I will here call *presentation space* … *presentation space* refers to how much can be presented (displayed, visualized) at a given time to the dictionary user. (Lew, 2014)

Lew argues that the visualized dictionary articles should not be too long and this argumentation is, as a rule, correct because it constitutes a first step to avoid information overload. However, this only gives rise to other questions: How much data should be displayed at a time? How can the specific types and exact amount of data be determined? How can it be adapted to screens of various sizes, from big computer screens over the ones used in laptops, tablets and mini-iPads to the various types of screens used in mobile phones? And how can lexicography take into account the growing user demand for still more personalised products?

> The demand among consumers generally is for products that match their individual needs more precisely – an expectation that is already transforming businesses like television and popular music, for example. In dictionary terms, this implies both customization and personalization. (Rundell, 2010: 172)

The big challenge facing lexicography is how to fulfil the transformation of dictionaries and obtain what Rundell calls *customization* and *personalization* and Tarp (2011) *individualization*. This requires undoubtedly highly efficient methods. Although

there are several opinions about the relevance of lexicographical theory, as a starting point it is difficult to see how such methods can be developed without input and guidance from an advanced theory, for instance, the *function theory* upon which this contribution is based (cf. Bergenholtz & Tarp, 2003; Tarp 2008; Fuertes-Olivera & Tarp, 2014).

4. The core of lexicography

So, what prevents lexicography from taking the great leap into the Brave New e-World? The sad answer seems to be business as usual, expressed in ingrained habits, conservatism, stubbornness and lack of visions. Old habits already problematic in the printed world have been transferred to the virtual universe. Two postulates have been repeated so often that they have become a sort of mantra among many lexicographers:

- Dictionaries are compiled in order to describe the language.
- All decisions taken by lexicographers are linguistic decisions.

Rundell (2012: 60), for instance, defines "the core task for lexicographers" as "analysing the evidence of language in use in order to identify what is likely to be relevant to dictionary users". This statement is, in my understanding, highly problematic. The analysis of "the evidence of language" is, of course, a *sine qua none* when preparing quite a number of specific data categories and dictionaries, especially general ones, but it can never be the core task for lexicography as such because there are a big number of other data categories and dictionaries which do not presuppose any analysis of the evidence of language. Let history and practice speak for themselves.

One of the earliest specialised dictionaries published in Britain was John Harris' *Lexicon Technicum* with the eloquent subtitle "*an Universal English Dictionary of Arts and Sciences: Explaining Not Only the Terms of Art, but the Arts Themselves*". In his Preface, the author explains in no uncertain terms that his dictionary does not only treat "the evidence of language":

> That which I have aimed at, is to make it a Dictionary, not only of bare Words but Things; and that the Reader may not only find here an Explication of the Technical Words, or the Terms of Art made use of in all the Liberal Sciences, and such as border nearly upon them, but also those Arts themselves; and especially such, and such Parts of them as are most useful and advantageous to Mankind. (Harris, 1704: Preface)

Another early bird among British specialised dictionaries is Malachy Postlethwayt's *Universal Dictionary of Trade and Commerce*, published in various editions in the mid-eighteenth century. The author was a British economist and publicist of certain renown during this époque; the dictionary was originally a translation and adaptation of a similar French dictionary, Jacques Savary des Bruslons' *Dictionnaire universel de commerce*, published first in 1723 and subsequently in various editions and translations. Postlethwayt's dictionary is interesting for various reasons, one of them being that the

author two years before its first appearance published a small essay where he explained his motives for this excursion into the world of lexicography (Postlethwayt, 1749). Here he wrote that the objective of the planned dictionary was to "more particularly accommodate the Fame to the Trade and Navigation of the British Empire"; this was done in the light of a serious problem he had observed, namely, that the relevant people frequently did not have a "satisfactory knowledge of Facts in complicated matters of a commercial nature", and that, in addition, these people had neither time nor the possibility to obtain this knowledge from the existing sources (Postlethwayt, 1749: 2). He argued:

> Foreign and domestic trade admitting of so infinite variety of matter, and the knowledge communicated to the world, by those skilled and experienced therein, being scattered in an infinity of volumes, it is no easy matter to have immediate recourse to what may be occasionally requisite… A subject of this extensive nature therefore being reduced to the form of a Dictionary, for alphabetical reference, seems the most naturally adapted to answer these desirable purposes, and especially so, as the compilers can have no motive to deceive. (Postlethwayt, 1749: 2)

As an example of the practicability of this idea, he referred explicitly to Savary des Bruslons' *Dictionnaire universel de commerce*, a "celebrated work" which had proven "how far an universal knowledge of commerce is capable of being reduced into the like form" (Postlethwayt, 1749: 3). It was these ideas that he himself applied shortly afterwards in his own dictionary. If we bear all this in mind, it seems nothing short of impossible to meaningfully characterise Postlethwayt's *Universal Dictionary of Trade and Commerce* and Savary des Bruslons' *Dictionnaire universel de commerce* as practical results of "analysing the evidence of language". Both authors were subject-field experts within trade, commerce and economics, and their respective dictionaries were compiled as a result of an extensive study of the *evidence of trade*, *commerce and economics* as developed in their life time.

Many similar examples can be found in the history of lexicography. For instance, in a highly interesting article on the development of economic dictionaries and encyclopaedias in the Iberian Peninsula, Astigarraga, Zabalza & Almodovar (2001), three specialists in the history of economics, write the following about two of these dictionaries and their author:

> José Canga Argüelles, a liberal economist and the then Minister of public finance, was the first Spanish economist to embrace the concept. In 1818, in the middle of the Absolutist period, he had been refused permission to publish his *Diccionario económico, estadístico y de hacienda...* but finally the project saw the light of day under the title of *Diccionario de hacienda para el uso de los encargados de la suprema dirección de ella* (1826-1827). He followed this up with the *Diccionario de hacienda, con aplicación a España* (1833-1834). There is no doubt that these two dictionaries contained the most worthwhile articles on economics written by a Spanish economist during the first half of the 19th century. Both works... were intended for use in training senior civil servants in the Spanish Department of public finance... (Astigarraga Zabalza & Almodovar, 2001: 29).

In his Preface to both dictionaries, the author himself confirms this objective of his dictionaries which he describes as a "small library of public finance":

> … el presente diccionario se puede mirar como una pequeña biblioteca de Hacienda. No se hallan en ella largas disertaciones que estarían en contradicción con el estilo de esta clase de obra, ni tampoco las ordenanzas y reglamentos dados para el manejo de las rentas. Generalizar los conocimientos de la Ciencia de Hacienda en todas sus relaciones, facilitando noticias, aunque poco comunes, precisas á los encargados de su dirección, son los objetos que me propuse al emprender la presente obra. (Canga Argüelles, 1826: vii)

This approach to lexicography is not only of historical relevance, but also very relevant in the twenty-first century. For instance, in a recent book which analyses how the concepts of *crisis* and *business cycles* have been treated in economic dictionaries and encyclopaedias during the past centuries, Daniele Besomi has studied more than 650 works published since 1709. The Swiss researcher and specialist in the history of economics concludes the following about their role in society:

> These reference works played an important role in the popularization but also in the systematization of knowledge during the nineteenth century and the early twentieth century, and – judging from the continuing exponential increase in their publication – are still widely used in the support of teaching and, to a lesser extent, research. What is recorded in dictionaries is therefore rather influential, in particular for those works recognized by contemporaries to be authoritative. (Besomi, 2011: 4)

So, what are the conclusions which can be drawn from this lexicographical reality? Some lexicographers may claim that the mentioned dictionaries are not dictionaries because they are not like the ones they themselves make, but such a conclusion would represent a very narrow approach to their own discipline and lexicographical practice in general. Others may claim that these dictionaries are only a few and therefore not relevant to lexicography, but this would contradict the fact that specialised dictionaries of various kinds clearly seem to outnumber their general-language relatives in terms of titles:

> In the present-day library, scientific, technical, and other specialized dictionaries greatly outnumber all other kinds, and present special problems for the lexicographer and historian alike. (Hoare, 2009: 47)

All this points to another conclusion which does not ignore present and past lexicographical reality. If all these very different works can be grouped under the category of dictionaries (or similar names), then there must be *something common to all of them*, and this something cannot be linguistics nor any of the other hundreds of disciplines interacting with lexicography in the compilation of dictionaries of whichever type. These uniting elements and aspects common to all types of dictionary can be considered the *real core of lexicography*. If an abstraction is made from all the specific

characteristics in terms of language, topic, users, purpose, media, etc., then the fundamental uniting elements and aspects can be described as follows:

- Dictionaries are utility *tools*
- designed for *consultation*
- and produced with the genuine purpose of meeting *punctual information needs*,
- which specific types of potential *user*
- may have in specific types of extra-lexicographic *situation*,
- by providing access to carefully prepared *data*
- from which the users can retrieve *information*
- which can subsequently be used for *different purposes*.

The above points constitute the basic axiomatic statements of the lexicographical function theory. All human-made tools are produced with a genuine purpose, namely to satisfy different types of human needs. Consequently, immediately after defining dictionaries as utility tools, the function theory establishes the specific category of *human need* (information) to be met by these tools, the *means* (lexicographical data) to meet this category of need as well as *how to use the tools* (consultation).The binomial global-punctual does not represent a relation between big and small, or more or less, but between the whole and the part, i.e. a relation determined mainly by quality, and not only by quantity. This distinction is fundamental for the solution of the present challenges of e-lexicography. Finally, the theory clarifies that the punctual information needs are always closely related to users with specific (lexicographically relevant) characteristics as well as to the type of situation where they occur. The underlying idea is that different types of user may have different needs, just like one and the same user may have different needs in different situations. In this respect, the function theory distinguishes between four fundamental extra-lexicographical situations, each of which can be further sub-divided:

1. *Communicative situations* where a need to solve a communication problem may occur (e.g. text reception and production in the mother tongue or a foreign language, translation from and into the mother tongue, and text revision).
2. *Cognitive situations* where a need for knowledge about any subject may occur.
3. *Operative situations* where a need for instructions on how to perform a physical or mental action may occur.
4. *Interpretive situations* where a need to interpret and understand a non-linguistic sign, signal, symbol etc. may occur.

Most dictionaries are designed to solve needs related to the two first situations, but there are also numerous examples of dictionaries dealing with the two latter. The extra-lexicographical situations are not only crucial when defining the types of information need to be satisfied, but also when determining which of the foreseen users' many characteristics are relevant in each dictionary project. Spanish mother-tongue speakers'

proficiency level in English is, for instance, totally irrelevant when designing a monolingual Spanish dictionary for text production, but it is highly relevant when preparing a dictionary for Spanish learners of English as a foreign language. So, when determining the information needs to be met by a concrete dictionary, first it is necessary to establish the situation(s) to be covered and then the relevant user characteristics.

All this leads to a much broader vision of lexicography than the one underpinning the statement by Rundell (2012) quoted above. In this broader vision, "the core task for lexicographers", or for the lexicographical team as a whole, is four-fold: 1) to establish the exact types of information need for each type of foreseen user in each type of situation to be covered by the dictionary in question; 2) to determine, upon this basis, the exact types of data required to meet these needs; 3) to prepare the corresponding lexicographical data by means of the most adequate methods; and 4) to ensure that these data can be accessed as quickly and easily as possible by the foreseen users.

5. Theory and methodology

A general theory of lexicography should necessarily include *all aspects* of lexicography (including *all types* of dictionaries and other lexicographical works) and *cannot be restricted only to one subset of dictionaries*, e.g. dictionaries where special knowledge of linguistics has been required. Such a general theory cannot ignore that lexicographical works are *multi-faceted* cultural artefacts and utility tools which, during the millennia, have met *a wide range of different needs* detected in society and covered almost all spheres of human activity and knowledge. It should not take its point of departure in the differences that separate all these works in terms of their specific content, structure etc., but in the aspects and elements that are common to all of them. The uniting core elements, which have been discussed above, are the reason why supporters of the function theory consider lexicography to be an independent discipline with a big interdisciplinary vocation, i.e. independent but not isolated from the "outside world".

As an independent discipline, lexicography must develop its own system of theories, methods, techniques, etc. It cannot blindly take over theories, methods, techniques, etc., from other disciplines without submitting them to critical analysis with a view to determining what should be rejected, what can be used as it is, and what can be used only after being adapted to suit the particular nature of lexicography. This is a *fundamental principle of methodology* applied by the function theory in the interaction between lexicography and other relevant disciplines, cf. Tarp (2014). There are numerous examples of how this principle has been implemented. For instance, methods developed within linguistics in order to analyse, describe or normalise language may be perfectly adequate for this purpose. However, they are frequently not sufficiently focussed on the specific tasks to be solved by lexicography and are, therefore, not necessarily the most appropriate when it is a question of solving the concrete types of information need experienced by different types of user in communicative situations.

With this in view, Bergenholtz (2003) has argued that the traditional linguistic and language-political methods of description and prescription need to be accompanied and, in a certain sense, substituted by another method within lexicography, i.e. proscription (recommendation), especially in relation to dictionaries designed to assist users with text-production problems. Similarly, Bergenholtz & Agerbo (2014a) have shown that the linguistic concept of meaning cannot be copied and applied in dictionaries for text reception and production. Tarp (2008) has argued that the traditional linguistic distinction between semantic and encyclopaedic knowledge is completely irrelevant when writing lexicographical definitions. Gouws (2013) has discussed why lexicography cannot uncritically take over the synonymy concept used by linguistics. Tarp (2009) has done the same with regard to the homonymy and polysemy concepts of linguistics. Bergenholtz & Gouws (2014) have argued that lexicography needs a classification of multiword combinations different from the one normally used within linguistics. Tarp (2013) has illustrated how the translation model needed to guide the design of translation dictionaries is different from the models used in translation science, though inspired by these. Bothma & Tarp (2014) have shown that relevance theory, as it has been developed within information science, needs to be complemented with an additional dimension related to the functionality of the tool when applied within lexicography.

It is important to stress that the application of this important principle of function-driven methodology does not represent a rejection of linguistics or any other discipline with which lexicography interacts. It simply aims at regulating the relation between lexicography and these disciplines in the sense that the former should always be in command in order to adapt everything to its specific requirements and, thus, guarantee maximum user satisfaction.

The methods used to prepare data categories in general dictionaries are frequently not identical if these dictionaries have different functions, and neither are the methods used in general and specialised lexicography. Bergenholtz & Agerbo (2014b), for instance, provide a detailed description of a specific method developed with a view to identifying, selecting and distributing meaning elements in a general monolingual dictionary designed to assist users with text-reception problems. By contrast, Fuertes-Olivera (2014) shows how a completely different method is required writing definitions in specialised dictionaries. While the two former authors use text corpora to extract and identify meaning elements, the latter relies heavily on subject-field experts. Similarly, Fuertes-Olivera (2013) rejects the value of text corpora for lemma selection in specialised dictionaries for translation and recommends instead another method. Other differences between methods used in general and specialised dictionaries with different functions could be mentioned.

6. How to detect user needs

When a lexicographer, or a publishing house, starts a new dictionary project, the first thing to do is to determine the concept, i.e. the overall design, of the dictionary and this should, in all cases, be based upon a determination of the types of user needs to be attended. How can this be done? Fuertes-Olivera & Tarp (2014) have argued that there are at least four known ways of doing this within lexicography:

1. Business as usual
2. Personal knowledge
3. User research
4. Deduction

None of these methods are perfect, but one is more realistic than the others. The two authors argue that the first method frequently results in low-quality products without being able to guarantee the needed innovation; that the second may be adequate for certain dictionaries but is too artisanal to solve the complex problems in the current transition to e-lexicography; and that the third method, when based upon scientific principles, may lead to detection of the relevant needs but that it is too costly and time-consuming to be used in each and every dictionary project with its own specific characteristics. They therefore recommend the functional approach.

As mentioned above, the needs which can be met by lexicographical works are not abstract but very concrete needs which – apart from the specific topic in question – depend on the situation(s) where they occur and the relevant characteristics of the users. With this in view, the method to determine the relevant needs is the following: one or several experts, who have a profound knowledge of the topic, situation(s) and foreseen users, make use of their experience in order to deduce the needs to be covered by the dictionary. This is always done under the guidance of a lexicographer.

Let us take the example of a teacher of business communication in a foreign language. If such a person has marked thousands of exercises and essays during the years and read the students' protocols related to this activity, and if this person thereafter has discussed the corresponding problems with the students, inclusive those related to information search and the use of reference works, then he or she will undoubtedly possess a profound knowledge of the problems and needs of this particular group of potential users in terms of foreign-language business communication. Hence, if this person works together with an expert trained in lexicographical theory and practice, then it would be perfectly possible for the two of them together to deduce and typologise these needs and determine which of them are lexicographically relevant. The thousands of marked exercises, essays, protocols and subsequent discussions with students will most often constitute a much better empirical basis than the results of a few dozen observations or the dubious data provided by hundreds of questionnaires.

The whole process can normally be carried out within a few hours. As such, the method is relatively easy and quick to apply. Until now, it has shown very good and

promising results materialised in a big number of dictionaries, both general and specialised, monolingual and bilingual, produced by supporters of the function theory. The method may not be perfect but it is capable of determining the huge majority of relevant user needs, even some occurring only very seldom. The results can easily compete with those obtained by user research but using only a fraction of the time and resources required to base the dictionary concept on such research in each case.

Once the lexicographers have elaborated a detailed list of user needs it is relatively easy to establish the *data categories* which are required to cover these needs as well as the relation between them. However, it is not enough to decide that the dictionary should contain inflection, definitions, collocations, etc. It should also be clarified how these categories should be understood and treated in each case. For instance, should the lexicographers apply the method of proscription, prescription or description when dealing with orthography, inflection or syntax? How should meaning be presented and at what expertise level? What types of collocation, synonym and antonym concepts should be used? In this respect, a large number of academic articles have been published showing how the various categories could be treated in specific dictionaries, some of them mentioned in the previous section.

7. Classification of online dictionaries

Online dictionaries can be classified in five main types according to the technology applied and the final result presented to the users. These five types are: 1) *Copycats*, 2) *Faster Horses*, 3) *Stray Bullets*, 4) *Model T Fords*, and 5) *Rolls Royces*, cf. Fuertes-Olivera & Tarp (2014: 13-18). This exoctic terminology is based on a quotation from Henry Ford who, when introducing his famous Model T Ford, was asked if he had consulted his users before inventing this model. According to the legend, his laconic answer was that if he had asked the users what they wanted, they would have said "faster horses".

The lexicographical *Copycats* are exact copies of existing dictionaries which have been placed on the Internet as PDF files or photos. There are two kinds, historical dictionaries like the ones quoted above and completely new dictionaries. The *Faster Horses* are dictionaries, whether new ones or already existing printed ones, which have been published online making use of various search options in order to provide quick data access, but where the articles visualised are moulded after traditional *static articles* in printed dictionaries with no variation from consultation to consultation. The majority of existing online dictionaries are either Copycats or Faster Horses. All these dictionaries are basically paper or paper-like dictionaries placed on the Internet without applying the available technology in order to improve the functionality of the lexicographical product. By contrast, the *Stray Bullets* have made extensive use of the new technology, but in this case technology has taken the command and sidelined important lexicographical principles in terms of quick and easy data access and information retrieval at the expense of user satisfaction.

As a result, only a few online dictionaries have so far taken the big step into the Brave New e-World. These dictionaries are represented by the *Model T Fords* and are characterised by *dynamic articles* where the data appearing on the screen is adapted both to the type of *user* and the type of *situation* where the needs occur. An example of this type of online dictionary are the *Diccionarios de Contabilidad,* a series of interconnected Spanish, Spanish–English, and English–Spanish dictionaries with various functions and search options. The fifth type of online dictionary is the *Rolls Royce* which is characterised by *dynamic*, *individualised articles and data* adapted to the user's concrete needs *in each consultation*. This type of online dictionary still constitutes an empty category in the sense that no dictionary has yet qualified to this title although there are a small number of dictionaries trying to go beyond the Model T Ford. The present challenge of online lexicography is to completely leave the printed media behind and start the generalised production of lexicographical Model T Fords with a view to take the future step towards the further individualisation in the form of lexicographical Rolls Royces.

8. Principles of online dictionaries

Information overload – or data overload – is not just a problem in terms of quantity, but also of quality. There is no absolute criterion to determine when there is a case of data overload in a dictionary article. It all depends on the user's exact needs as well as the size of the screen where the data are displayed. If there are more data than required to meet these needs, then it is a clear case of *absolute overload*. Similarly, if there are more data than can be visualised simultaneously without scrolling down, or than the predicted type of user can be expected to overview, then it may be a case of *relative overload*, even if the data displayed are all relevant; this may, for instance, occur in dictionaries with communicative functions where users most often demand a quick answer to their problem. Absolute data overload relates both to the needs which a specific type of user may have in a specific type of situation (*functional overload*), and to the needs which a concrete, individual user may have in a concrete situation (*concrete overload*). The solution to functional overload is a lexicographical Model T Ford, whereas the solution to concrete overload is a future lexicographical Rolls Royce. As to relative data overload, the solution, or at least mitigation, of this problem requires a number a special techniques, some of which will be briefly treated in Section 9.

Apart from avoiding data overload, there are other important principles from which the production of future online dictionaries would benefit considerably:

1. Users should be able to access the data required in each consultation *as quickly as possible*.
2. The database should include *as much data as possible*, i.e. as much data as possible relevant to the type(s) of dictionary in question.

3. The specific dictionary should be able to present *as much data as possible* in terms of all possible consultations, i.e. the entire body of hypothetic articles resulting from these consultations.
4. The individual articles, namely, the dynamic data presented on the screen in each consultation, should include *as little data as possible*, i.e. exactly the types (qualitative criterion) and amount (quantitative criterion) of data needed by the user in each situation (Model T Fords) or each consultation (Rolls Royces), neither more nor less.

The underlying idea here is that a necessary distinction should be made between the *database* which stores all the lexicographical data; the *dictionary as such* which should be understood as the totality of lexicographical data (articles) that may be displayed on the screen in the totality of possible consultations; and the *individual articles* which are represented by the lexicographical data displayed on the screen in each consultation. Online and other electronic dictionaries are not databases, but consultation tools based upon databases from which they take in the data required to meet their users' information needs; as such, they may even feed various dictionaries at the same time, cf. Bergenholtz & Bergenholtz (2013).

It goes without saying that a good and well-composed database should contain as much relevant data as possible, and so should the dictionary in question. This is absolutely necessary if the dictionary should be capable of solving the big variety of the user needs in all possible consultations. However, if information overload should be avoided and quick and easy data access and information retrieval guaranteed, then it is just as necessary that the individual articles contain as little data as possible. This principle is of special importance in connection with small handheld devices such as cell phones and tablets. With the study of relevant user characteristics and the social contexts where lexicographically relevant needs may occur, lexicographers have a powerful weapon to determine the types and amount of data to be presented to the users in each consultation. However, in order to achieve this goal it is also necessary to apply a number of special techniques, of which some of the most important will be discussed in the following Section.

9. Techniques used in online dictionaries

There are already various available techniques which, correctly applied, can be used to produce lexicographical Model T Fords and take the first steps towards the future Rolls Royces, i.e. towards the individualisation of user satisfaction. In the following, we will briefly look at some of these techniques based upon the reflections made by an information scientist (Bothma, 2011):

Filtering. The easiest way to produce a lexicographical Model T Ford is to design it with only one function and therefore only provide the data required to assist one type of user in one type of situation. However, most often the overall dictionary project is designed to cover various functions. In such cases, the user interface may include

interactive options where the users can define themselves and the situation where they experience problems or needs (see Example 5). In accordance with the resulting user profile and situation, the pre-programmed system will then filter the data contained in the database and automatically calculate the types and amount of lexicographical data required to fulfil the needs occurring for this specific user type in the situation in question. These data will be the only ones presented on the screen in the concrete consultation. While the user profile can be made once and for all and only needs to be refined when the user's relevant characteristics change, the description of the situation has to be supplied to the system when starting each new task. Although it should always be possible to "re-saddle" in the middle of the process, the important thing is that the user does not need to go through all these time-consuming steps for each consultation. A precondition for a successful application of this and some of the following techniques is that each single data has its own, individual number in the database and that various data are not mixed together.

Adaptive presentation. There are various techniques in terms of adaptive hypermedia but here we will mention only pop-up windows activated by either mouse-moving or clicking. Very often the lexicographer decides that not all data needed in a specific consultation should be presented at first sight. There could be various reasons for this, all of them related to **relative** data overload as defined in the previous section. Whatever the reason, the lexicographer can choose to "hide" some data, which can easily be accessed by moving the mouse over a specific word or area, or clicking on a link. Then a pop-up window will appear with more data. It could, for instance, be additional collocations, an illustration, or a table with the inflected forms of a Spanish verb, which would occupy a large amount of screen space (see Example 10). In this respect, adaptive presentation is not only a technique to provide more data of the already existing types (*quantitative adaption*) but also new data types (*qualitative adaption*). This technique is especially relevant when users are working with small screens like tablets, mobile phones or other hand-held devices, cf. Kwary (2013).

Indexing. A technique very similar to the above one is indexing which is mainly used when the article providing the needed data is very long, i.e. especially in dictionaries with cognitive functions. In these cases, the articles visualised start with a preliminary index which can be further expanded by means of hypermedia (like in *Wiktionary*). The application of this technique immediately provides the user with an overview of the content. In addition, the possibility of expanding an index by means of hypermedia allows the user to go directly to the section of the article which is pertinent in each case.

Article modelling. Another technique is that each individual user of an online dictionary will be given the option to design his or her own master article in terms of the types of data wanted and their arrangement on the screen. The system will then automatically filter the available data and present them as indicated by the user. An online dictionary allowing for article modelling in terms of data types is the Swedish *Lexin* (see Example 4).The design of a master article may be accomplished when the users enter the online dictionary for the first time, but it can also be done when they

begin an extra-lexicographical activity, start a consultation, and even in the middle of a consultation. In this way, it is possible to re-saddle whenever necessary in order to individualize the final data presentation. In order to illustrate the degree of individualisation which can be obtained with this technique, Tarp (2012: 260) has taken the hypothetic example of a relatively simple online dictionary with 10 data types addressed to each lemma. If such a dictionary was designed for fully individualised access and the users was given the possibility to choose only those data that they wanted to visualise, then 1,023 (2^{10}-1) possible data combinations could be displayed on the screen for each lemma. And to this should be added that the amount of possible combinations in each individual article would rise to the gigantic sum of 4,037,913 (1!+2!+3!+4!+5!+6!+7!+8!+9!+10!), i.e. more than four million combinations, if the users were also allowed to arrange the data in the order desired by each of them individually.

Example 4: Pop-up window from Lexin showing options allowing the users to design individual master articles by including and excluding data types

Annotation. In the Web 2.0 environment, users are allowed to add their own data to existing documents without changing the original. When this technique is employed, users can write comments, incorporate supplementary data and recommend the use of certain terms. In this way, annotation may contribute to the future individualisation of lexicographical works and may above all be useful in the framework of predefined user groups such as companies, branches, public entities, research groups, etc.

Reuse of data. An online dictionary connected to a database may not always contain enough data to fulfil the user's information needs in a specific consultation. In such cases, it could be advantageous to provide access to additional data stored in external sources such as corpora, databases, and the Internet in general. The data in question could be texts providing supplementary background information, additional collocations, example sentences showing the usage of collocations and other linguistic features, or so-called contextual definitions found in existing texts. This kind of data reuse can be achieved by various techniques. One of these are links to manually selected external texts and data but this technique "requires a tremendous input in terms of time and effort from the lexicographer" (Bothma, 2011: 92). Another techniques is to link directly to external databases with different types of content, for instance "general and specialized corpora, allowing users to search for examples of any word, pattern, or linguistic feature they are interested in" (Rundell, 2010: 174). According to Heid, Prinsloo & Bothma (2012):

> Electronic dictionaries should supply a natural bridge for the user from the dictionary article to a variety of internet resources to enhance access to potentially relevant information. In total, the user should be able to find a wealth of digestible information without being overloaded. (Heid, Prinsloo & Bothma, 2012: 270)

In this respect, one of the visions of online lexicography today is not only to reuse external data but also to *repackage* them by adapting them to the specific information needs of the user in each situation. Some initial steps have been taken in this direction but generally it can be said that the required technique, which involves complex validation problems, still has to be developed and perfected in order to convince.

10. An example

In this section we will discuss an online Business Dictionary which, in spite of its title, aims at innovation instead of "business as usual". The dictionary, produced in international cooperation between lexicographers from Denmark, Cuba, Spain and South Africa, is still in the construction phase and is expected to be accessible on the Internet in 2016. Its main purpose is to provide assistance to English and Spanish-speaking professionals, business people, civil servants, secretaries, translators, students and other potential users engaged in business and other type of professional communication. The situations to be covered by this e-tool are L2 text production, L1-L2 translation, L2 text reception, L2-L1 translation, and L1 text production for users with either Spanish or English as their mother tongue or first language. This makes a total of 10 different lexicographical functions, to each of which corresponds a master article with specific data categories. In order to provide access to the most appropriate type of article, the planned user interface will contain inter-active options where the users, with only one click, can define themselves together with the type of communicative situation in which they need assistance (see Example 5).

Business Dictionary
– Diccionario de Comunicación Empresarial

Enter a Word...	SEARCH

Haga clic en su problema para ser mejor atendido
Mi lengua materna es español y necesito ayuda para

- ❑ escribir un texto en inglés
- ❑ traducir un texto del español al inglés
- ❑ comprender un texto en inglés
- ❑ traducir un texto del inglés al español
- ❑ escribir un texto en español

Click on your problem in order to get better assistance
My first language is English and I need assistance

- ❑ to write a Spanish text
- ❑ to translate from English into Spanish
- ❑ to understand a Spanish text
- ❑ to translate from Spanish into English
- ❑ to write an English text

Example 5: User interface with interactive options allowing for data filtering

The Business Dictionary is not designed to assist L1 text reception in general as the foreseen target users are supposed to understand most of the L1 words and collocations included in the dictionary. Only a relatively small number of terms considered to be difficult will be defined. The general writing of definitions would considerably delay the compilation process and the moment where the dictionary can be published; it is, therefore, postponed to a later moment if it appears to be necessary. Besides, the main aim of the dictionary is to solve problems related to L2 communication where text reception and production directly in L2 seem to be the two most frequent situations today, rather than translation of full texts in both language directions. The overall concept is based on the function theory as well as a *lexicographical* study of the respective situations and the corresponding phases and sub-phases where potential users may experience *lexicographically relevant* needs (see, for instance, Tarp, 2004, 2013).

If a Spanish first-language speaker is writing a text directly in English (i.e. without translating a previous text or outline in Spanish) and encounter any kind of problem related to the use of an English word, he or she can simply enter this word in the search field and simultaneously click on *escribir un texto en inglés*. The concrete article appearing on the screen will then be based on a master article containing the following data categories:

English lemma (uk/us) part of speech
<Inflection>
Abbreviated form / full form
Definition

<u>Link to external source</u>
(meaning discrimination) **Spanish equivalent**
(meaning discrimination) **Spanish equivalent**
etc.
syntactic properties
mini-rule: Example sentence in English. Example sentence translated into Spanish. (Note written in Spanish)
mini-rule…
etc.
collocations
English collocation (uk/us) Spanish translation. (Note written in Spanish)
English collocation...
etc....
synonyms
<u>English synonym</u>
etc...
antonyms
<u>English antonym</u>
etc...
<p align="center">Example 6: Article assisting a Spanish user writing an English text</p>

Apart from the *English lemma*, *part of speech*, and at least one *Spanish equivalent* (supplied to confirm the meaning of the lemma), the other data categories contained in this master article are optional depending on the characteristics of each lemma. Unfortunately, the limits of this contribution do not permit a thorough description of these categories and the argumentation for the inclusion of each of them. Here it should only be mentioned that all words underlined represent direct links to the corresponding articles.

If the Spanish user instead clicks on *traducir un texto del español al inglés*, an article based upon the following master article will be displayed:

Indication of Spanish search word(s)

(meaning discrimination) <u>**English equivalent**</u> (uk/us) part of speech
(meaning discrimination) <u>**English equivalent**</u> (uk/us) part of speech
(meaning discrimination) <u>**English equivalent**</u> (uk/us) part of speech
(meaning discrimination) <u>**English equivalent**</u> (uk/us) part of speech
etc...

Example 7: In-between page leading to an English article designed to assist a Spanish user translating a Spanish text into English

The master article shown in Example 7 is conceived as an in-between article where the user, if required, can click on one of the English equivalents and be directed to the corresponding English article (as the one shown in Example 6). The underlying philosophy is twofold: 1) if a user just needs a reminder and can solve the problem with

only an English equivalent, then there is no reason to provide a lot of data irrelevant to this user; and 2) if the user do need additional information on the equivalent (inflection, syntactic properties, collocations, etc.) in order to produce a correct English text, then the article would be completely overloaded if the corresponding data were attached to each of the English equivalents furnished (in many cases more than five).

If a Spanish user clicks on either *comprender un texto en inglés* or *traducir un texto del inglés al español*, the article appearing on the screen would be based upon one of the following master articles, respectively:

English lemma (uk/us) part of speech
<Inflection>
Abbreviated form / full form
Definition
Link to external source
(meaning discrimination) **Spanish equivalent**
(meaning discrimination) **Spanish equivalent**
(meaning discrimination) **Spanish equivalent**
(meaning discrimination) **Spanish equivalent**
etc...
collocations (to unfold)
Example 8: Article assisting a Spanish user understanding an English text
English lemma (uk/us) part of speech
<Inflection>
Abbreviated form / full form
Definition
Link to external source
(meaning discrimination) **Spanish equivalent**
(meaning discrimination) **Spanish equivalent**
(meaning discrimination) **Spanish equivalent**
etc...
collocations
English collocation (uk/us) Spanish translation. (Note written in Spanish)
English collocation (uk/us) Spanish translation. (Note written in Spanish)
English collocation (uk/us)…
etc....
Example 9: Article assisting a Spanish user translating an English text into Spanish

The two master articles provided in Example 8 and 9 are in many aspects similar. One could with reason discuss whether English collocations and their Spanish translation should also be provided immediately in the article assisting English text reception (Example 8), but here it has been decided to include them as a unfold option in order to primarily focus on the cumulative Spanish equivalents. (This focus will be stressed by the fact that the remaining categories, i.e. *abbreviated form*, *full form*, *definition* and *link to external source*, will only be displayed in relatively few concrete articles). In this respect, the mayor difference between the two master articles is that problems related to L2-L1 translation do not necessarily stop with the provision of a L1

equivalent, as the **user**s may also experience some production problems in their first language, for instance, in relation to the use of the right collocations. Consequently, the article shown in Example 9 contains direct linking from the Spanish equivalent to the corresponding Spanish lemma.

Finally, if the Spanish-speaking user clicks on *escribir un texto en español*, any article visualised on his or her screen will be based on the following master article:

Spanish lemma part of speech
<Inflection> (to unfold)
Abbreviated form / full form
Definition
Link to external source
syntactic properties (to unfold)
collocations
Spanish collocation
Spanish collocation
etc...
synonyms
synonym
etc...
antonyms
antonym
etc...

Example 10: Article assisting a Spanish user writing a Spanish text

In terms of data categories, this master article is similar to the one presented in Example 6 with the exception of the categories related to the bilingual dimension in the latter. As the use of the right L1 collocations (as well as synonyms and antonyms) is considered to constitute the main problem for a Spanish-speaking user, it has been decided to focus on these data categories and provide the space-consuming inflection and syntax patterns as unfold options. However, if the users activate the latter they will get syntactic data based on the following mini-rule, here exemplified by the verb *rogar*:

mini-rule: Example sentence in Spanish. (Note written in Spanish)

rogar (que) + subjuntivo: Les rogamos nos manden su respuesta a la mayor brevedad posible. (Nota: "que" se omite normalmente en el lenguaje formal)

Example 11: Mini-rule in an article assisting a Spanish user writing a Spanish text

As can be seen, in the design of the "not business as usual" Business Dictionary it is foreseen that the information techniques *filtering*, *adaptive presentation*, and *reuse of external data by linking* will be applied. As such, the dictionary represents an example

of a typical lexicographical Model T Ford taking the first modest steps towards a more personalised tool with individualisation of user needs satisfaction.

11. Conclusion

The example of an on-going dictionary project presented in Section 10 contains the main conclusions of this article. Here, it is sufficient to quote Albert Einstein who in 1938, together with Leopold Infeld, wrote the following about the "real advance in science":

> 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. (Einstein & Infeld, 1938: 92)

Hopefully, lexicography will be inspired by these words from a one of biggest geniuses of the 20th century.

References

Dictionaries

BusinessDictionary.com = *BusinessDictionary.com*. Washington: WebFinance Inc. [Online: http://www.businessdictionary.com]
Business Dictionary = Sven Tarp (ed.): *Business Dictionary – Diccionario de Comunicación Empresarial*. (A series of 10 interconnected English, English-Spanish, Spanish, and Spanish-English dictionaries). (Under construction, expected to be available online at Lemma.com in 2016)
Diccionarios de Contabilidad = Pedro A. Fuertes Olivera, Pablo Gordo Gómez, Marta Niño Amo, Ángel de los Rios Rodicio, Ángeles Sastre Ruano, Sven Tarp, Marisol Velasco Sacristán, Sandro Nielsen, Lise Mourier and Henning Bergenholtz (2014): *Diccionarios de Contabilidad*. (A series of various interconnected Spanish, Spanish–English, and English–Spanish dictionaries). Hamburg: Lemma.com.

Diccionario de hacienda = José Canga Argüelles (1826-27): *Diccionario de hacienda para el uso de los encargados de la suprema dirección de ella*. Vol. 1-5. London: Imprenta española de M. Calero. [Online: books.google.com]

Diccionario de hacienda con aplicación a España = José Canga Argüelles (1833-1834): *Diccionario de hacienda con aplicación a España*. Vol. 1-2. Madrid: Imprenta de Marcelino Calero y Portocarrero. [Online: books.google.com]

Lexicon Technicum = John Harris (1704): *Lexicon Technicum: or, an Universal English Dictionary of Arts and Sciences: Explaining Not Only the Terms of Art, but the Arts Themselves.* London: D. Brown, J. Walthoe, J. Knapton, B. and S. Tooke, D. Midwinter, B. Cowse, T. Ward, E. Symon, E. Valentine, and J. Clark. [Online: books.google.com]

Lexin = Sven-Göran Malmgren (ed.)(2014): *Lexin*. Stockholm: Språkrådet. [Online version: http://lexin.nada.kth.se]

Oxford advanced learner's dictionary = Sally Wehmeier (ed.) (2010): *Oxford advanced learner's dictionary. Eighth edition*. Oxford: Oxford University Press.

Oxford Reference = *Oxford Reference*. Oxford: Oxford University Press. [Online: www.oxfordreference.com]

Universal Dictionary of Trade and Commerce = Malachy Postlethwayt (1774): *The Universal Dictionary of Trade and Commerce*. The Fourth Edition. London: W. Strahan, J. and F. Rivington, J. Hinton. [Online: archive.org]

Wiktionary = *Wiktionary. The free dictionary*. Wikimedia. [http://www.wiktionary.org]

Yongle Dadian = Jin Xie (ed.)(1408): *Yongle Dadian*. Vol. 1-11,095. China.

Other literature

Astigarraga, Jesús; Juan Zabalza and Antonio Almodovar (2001): "Dictionaries and Encyclopaedias on Political Economy in the Iberian Peninsula (18th, 19th and 20th Centuries)." *Storia del Pensiero Económico* 41: 25-63.

Bergenholtz, Henning (2003): "User-oriented understanding of descriptive, proscriptive and prescriptive lexicography." *Lexikos* 13: 65-80.

Bergenholtz, Henning and Heidi Agerbo (2014a): "Meaning identification and meaning selection for general language monolingual dictionaries." *Hermes. Journal of Language and Communication in Business* 52: 125-139.

Bergenholtz, Henning and Heidi Agerbo (2014b): "Extraction, selection and distribution of meaning elements for monolingual information tools." *Lexicographica* 30: 488-512.

Bergenholtz, Henning and Inger Bergenholtz (2013): "One database, four monofunctional dictionaries." *Hermes, Journal of Language and Communication in Business* 50: 119-125.

Bergenholtz, Henning and Rufus Gouws (2014): "A Lexicographical Perspective on the Classification of Multiword Combinations." *International Journal of Lexicography* 27(1): 1-24.

Bergenholtz, Henning and Sven Tarp (2003): "Two opposing theories: On H.E. Wiegand's recent discovery of lexicographic functions." *Hermes. Journal of Linguistics* 31: 171-196.

Bergenholtz, Henning and Sven Tarp (2010): "LSP Lexicography or terminography? The lexicographer's point of view." In P. A. Fuertes-Olivera, eds., *Specialised Dictionaries for Learners*. Berlin, New York: Walter de Gruyter, 27-37.

Besomi, Daniele (ed.)(2011): *Crises and Cycles in Economic Dictionaries and Encyclopaedias*. London, New York: Routledge.

Bothma, Theo J.D. (2011): "Filtering and adapting data and information in an online environment in response to user needs." In P. A. Fuertes-Olivera and H. Bergenholtz, eds., *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. London, New York: Continuum, 71-102.

Bothma, Theo J.D. and Sven Tarp (2014): "Why relevance theory is relevant for lexicography." *Lexicographica* 30: 350-378.

Canga Argüelles, José (1826): "Prólogo." En José Canga Argüelles: *Diccionario de hacienda para el uso de los encargados de la suprema dirección de ella*. Tomo 1. London: Imprenta española de M. Calero, iii-vii.

Einstein, Albert and Leopold Infeld (1938): *The evolution of physics*. New York: Simon & Schuster.

Fuertes-Olivera, Pedro A. (2013): "The Theory and Practice of Specialised Online Dictionaries for Translation." *Lexicographica* 29: 69-91.

Fuertes-Olivera, Pedro A. (2014): "Designing Online Dictionaries of Economics: Two Opposing Views." *Hermes. Journal of Language and Communication in Business*, 52: 25-40.

Fuertes-Olivera, Pedro A. and Sven Tarp (2014): *Theory and practice of specialised online dictionaries: Lexicography versus terminography*. Berlin, Boston: De Gruyter.

Gouws, Rufus H. (2013): "Contextual and Co-Textual Guidance Regarding Synonyms in General Bilingual Dictionaries." *International Journal of Lexicography*, 26(3): 346-361.

Harris, John (1704): "Preface." In J. Harris, ed., *Lexicon Technicum: or, an Universal English Dictionary of Arts and Sciences: Explaining Not Only the Terms of Art, but the Arts Themselves*. London: D. Brown, J. Walthoe, J. Knapton, B. and S. Tooke, D. Midwinter, B. Cowse, T. Ward, E. Symon, E. Valentine, and J. Clark.

Heid, Ulrich; D.J. Prinsloo and Theo J.D. Bothma (2012): "Dictionary and corpus data in a common portal: state of the art and requirements for the future." *Lexicographica* 28: 269-291.

Hoare, Michael Rand (2009): "Scientific and technical dictionaries." In A.P. Cowie, ed., *The Oxford History of English Lexicography*. Volume II: *Specialized Dictionaries*. Oxford: Oxford University Press, 47-93.

Kwary, Deny Arnos (2013): "Principles for the Design of Business Dictionaries on Mobile Applications." *Hermes, Journal of Language and Communication in Business* 50: 69-82.

Lew, Robert (2014): "Space Restrictions in Paper and Electronic Dictionaries and Their Implications for the Design of Production Dictionaries." In P. Bański and B. Wójtowicz, eds., *Issues in Modern Lexicography*. München: Lincom Europa. (Can be accessed on: www.staff.amu.edu.pl/~rlew/pub/Lew_space_restrictions_in_paper_and_electronic_diction aries.pdf)

Miller, George A. (1956): "The magical number seven, plus or minus two: some limits to our capacity for processing information." *Psychological Review* 63(2): 81-97.

Nichols, Wendalyn (2010): "I've have heard so much about you: Introducing the native-speaker lexicographer to the learner's dictionary." In I. J. Kernerman and P. Bogaards, eds., *English Learners' Dictionaries at the DSNA 2009*. Jerusalem: Kdictionaries, 29-43.

Nielsen, Sandro (2008): "The effect of lexicographical information costs on dictionary making and use." *Lexikos* 18: 170-189.

Postlethwayt, Malachy (1749): *A Dissertation on the Plan, Use, and Importance, of the Universal Dictionary of Trade and Commerce*. London: John and Paul Knapton.

Robins, Kevin and Frank Webster (1983): "The Mis-Information Society." *Universities Quarterly* 37(4): 344-355.

Rundell, Michael (2010): "What future for the learner's dictionary?" In I. J. Kernerman and P. Bogaards, eds., *English Learners' Dictionaries at the DSNA 2009*. Jerusalem: Kdictionaries, 169-175.

Rundell, Michael (2012): "'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical." In R. E. Vatvedt Fjeld and J.M. Torjusen, eds., *Proceedings of the 15th EURALEX International Congress*. Oslo: Institutt for lingvistiske og nordiske studier, 47-92.

Tarp, Sven (2004): "Reflections on dictionaries designed to assist the users with text production in a foreign language." *Lexikos* 14: 299-325.

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

Tarp, Sven (2009): "Homonymy and polysemy in a lexicographic perspective." *Zeitschrift für Anglistik und Amerikanistik* 57(3): 289-306.

Tarp, Sven (2010): "Functions of Specialised Learners' Dictionaries." In P. A. Fuertes-Olivera, eds., *Specialised Dictionaries for Learners*. Walter de Gruyter, 39-53.

Tarp, Sven (2011): "Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction." In P. A. Fuertes-Olivera and H. Bergenholtz, eds., *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. London, New York: Continuum, 54-70.

Tarp, Sven (2012): "Online dictionaries: today and tomorrow." *Lexicographica* 28: 251-265.

Tarp, Sven (2013): "What should we demand from an online dictionary for specialized translation?" *Lexicographica* 29: 146-162.

Tarp, Sven (2014): "Theory-Based Lexicographical Methods in a Functional Perspective: An Overview." *Lexicographic* 30: 58-76.

Toffler, Alvin (1970): *Future Shock*. New York: Random House.