

# From e-Lexicography to Electronic Lexicography. A Joint Review

Valeria Caruso, *Department of Literary, Linguistic and Comparative Studies, University of Naples 'L'Orientale', Naples, Italy*  
([vcaruso@unior.it](mailto:vcaruso@unior.it)/[vcaruso78@gmail.com](mailto:vcaruso78@gmail.com))

---

**Abstract:** Two recently published books outline the main issues of the current debate on lexicography. The first, *e-Lexicography* edited by Fuertes-Olivera and Bergenholtz in 2011, presents the standpoints of the lexicographical function theory on the future developments of dictionaries, while, in some chapters, current innovative tools are described, tools which allow customizations according to the user's type of need. The second volume, *Electronic Lexicography* edited by Granger and Paquot in 2012, presents different opposing views on what the dictionaries of the future will look like, such as the linguistic-oriented stance of Hanks and that of Tarp concerning theoretical lexicography. The dictionary projects that are described within these pages offer an interesting basis of comparison with those developed by the representatives of the function theory.

**Keywords:** COMPUTER ASSISTED LANGUAGE LEARNING (CALL), CORPUS LINGUISTICS, CUSTOMIZATION, DATABASES, DICTIONARY SURVEY, DICTIONARY USE, EFFICACY, EFFICIENCY, ELECTRONIC LEXICOGRAPHY, INFORMATION SCIENCE, LANGUAGES FOR SPECIAL PURPOSES, LEXICOGRAPHICAL FUNCTION THEORY, LINGUISTIC THEORY, MONOFUNCTIONAL DICTIONARY, P-DICTIONARIES, PRACTICAL LEXICOGRAPHY, USABILITY TESTING

**Opsomming: Vanaf e-leksikografie tot elektroniese leksikografie: 'n Gesamentlike beskouing.** Twee onlangs gepubliseerde boeke skets die belangrikste kwessies in die huidige debat oor die leksikografie. Die eerste, *e-Lexicography* in 2011 geredigeer deur Fuertes-Olivera en Bergenholtz, gee die standpunte van die leksikografiese funksieteorie oor die toekomstige ontwikkelinge van woordeboeke, terwyl, in sommige hoofstukke, huidige innoverende gereedskap beskryf word, gereedskap wat pasmaking volgens die gebruiker se soort behoefte toelaat. Die tweede boek, *Electronic Lexicography*, in 2012 geredigeer deur Granger en Paquot, bied verskillende opponerende standpunte oor hoe die woordeboeke van die toekomst sal lyk, soos die linguisties georiënteerde benadering van Hanks en dié van Tarp betreffende die teoretiese leksikografie. Die woordeboekprojekte wat in hierdie bladsye beskryf word, bied 'n interessante vergelykingsbasis met dié ontwikkel deur die verteenwoordigers van die funksieteorie.

**Sleutelwoorde:** REKENAARGESTEUNDE TAALANLEER (RGTA), KORPUSLINGUISTIEK, PASMAKING, DATABASISSE, WOORDEBOEKOORSIG, WOORDEBOEKGEBRUIK, DOELTREFFENDHEID, DOELMATIGHEID, ELEKTRONIESE LEKSIKOGRAFIE, INLIGTINGSWETENSKAP, TALE VIR SPESIALE DOELEINDES, LEKSIKOGRAFIESEFUNKSIETEOORIE, TAALTEORIE, MONOFUNKSIONELE WOORDEBOEK, P-WOORDEBOEKE, PRAKTIESE LEKSIKOGRAFIE, BRUIKBAARHEIDSTOETSING

## 1. Different points of view in the current debate

The short span of time within which two quasi-homonymous volumes have been published, mirrors the intensity of the debate upon the issues currently at stake in electronic lexicography. The first book, edited by Pedro Fuertes-Olivera and Henning Bergenholtz, appeared in 2011 and has a shorter title but an explicative subtitle: *e-Lexicography. The Internet, Digital Initiatives and Lexicography*; the second, entitled *Electronic Lexicography*, was published in 2012 and edited by Sylviane Granger and Magali Paquot. Both are miscellaneous volumes, even sharing some authors and co-authors – namely Fuertes-Olivera, Leroyer, Lew, Tarp and Verlinde – but the editorial projects and intended aims are different, and the reading of both is highly recommended to obtain an exhaustive picture of the topic.

*e-Lexicography*, by Fuertes-Olivera and Bergenholtz, is a celebratory volume on the occasion of the conferment of a Doctorate degree in Lexicography to Henning Bergenholtz, a highly symbolic event which goes some way towards the legitimatization of Lexicography as a separate and independent field among the academic disciplines, one of the crucial claims of the 'Aarhus School' of which Bergenholtz is one of the main representatives. The volume is a collection of the papers presented at the conference organized for Bergenholtz's Doctorate, an event, says the co-editor, Fuertes-Olivera, in his preface, during which many "new and provocative ideas" were put forth. The remark is not trivial, since some of the "ground-breaking" (Leroyer 2011: 125) tenets of the *lexicographical function theory* proposed by the Danish lexicographers have been met with disbelief by some esteemed lexicographers (Atkins and Rundell 2008) and metalexigraphers (Béjoint 2010).

The story is well-known but nevertheless worth a mention here. The researchers of the Aarhus University have proposed a new vision of what Lexicography should be about, considering it as a theoretical discipline within the Information Sciences, devoted to the study of all kinds of reference needs and tools, not limited by the genera, e.g. dictionaries, nor by the medium, e.g. paper or the Internet. There are actually "artefacts which, all things considered, should also be counted as members of the vast family of reference works, such as almanacs, atlases, catalogues, directories, guides, handbooks, reference manuals and so on" (Leroyer 2011: 124). This assumption is a consequence of the emphasis placed on the nature of dictionaries, conceived as tools that people refer to in order to satisfy their needs. Therefore every dictionary must be compiled not only with its user in mind, but also considering the situation in which it will be accessed. At the intersection between the user and the situation, the lexicographer is able to identify the functions that his/her dictionary must fulfil, e.g. spell-checking for school children, translational equivalences for a scholar preparing his paper in an L2, and so forth. Therefore the task to be fulfilled by the lexicographer is an effective data selection and presentation, minimizing user effort whilst satisfying information needs.

Though the premises of the *lexicographical function theory* on the practical

nature of the dictionary and its usefulness for users are uncontroversial, the broadening of the field of lexicography and the conferring of a theoretical status to it was questioned by those who remark on the 'practical' nature of the lexicographer's job (Atkins and Rundell 2008) which, in their view, is devoted to the "systematic", "internally-consistent" description of language, "driven by what the language data is telling us", with "subjective judgments [...] kept to a minimum" (Rundell 2012b: 48). In this respect the contribution of linguists and linguistic theory is paramount, and the computational analysis of language data is the most valuable revolution brought about by the digital text era, since the lexicographer's main concern is not the "synthesis" of data (Atkins 1993), i.e. selection and presentation, but rather mining from the raw corpora material.

Setting aside the provocative slogans and disbeliefs, it seems fair to say that these two parties are actually looking at different parts of the same object: the Aarhus School is committed to the data holder and data presentation, relying on the resources of Information Technology; the 'practical lexicographers' are concerned with data *per se*, and credit computational linguistics with the needed insight for effective language descriptions. However, for the users' benefit, it must be hoped that the research from both sides can combine rather than exclude each other, as seems to be suggested also by the wide-ranging volume collected by Granger and Paquot, hosting contributions from practical lexicographers and lexicography theoreticians as well.

Before dealing with this challenging topic, reviewing in more detail the chapters devoted to the description of different dictionary projects, an outline of the two volumes will be provided, reporting on all the points of view that the editors of the books have collected, since they are thought to be the key topics in the current electronic lexicography debate.

Nevertheless, a selection was necessary, and the less general themes will be dealt with only briefly, since the debate is chiefly concerned with three crucial issues: what features should dictionaries have in the future, what is lexicography, and which methodologies should it use? The empirical research on dictionary use is indeed the other challenging topic, particularly for a theory devoted to the functions of dictionaries, "in which user needs are key" (Granger 2012: 7). Actually, even if the perfect customizations that Tarp has in mind (see below) dismiss the question of profiling the needs of the target user, empirical findings on what can be considered 'ergonomic', should be collected.

With these three main issues in mind, it is possible after some preliminary remarks to outline the contents found in the books. While in the volume by Fuertes-Olivera and Bergenholtz, the reader can find not only the newest available dictionary formats, but also the details of future ones (in the chapters by Bothma, Spohr and Tarp), Granger and Paquot are more concerned with the overall current issues, giving a more varied inventory of specific sub-topics (translator's needs, 'alternative e-dictionaries' and software management tools, among many others) and visions, such as the lexicological perspective on the development of dictionaries described by Hanks.

In the following sections, the overall contents of each volume will be given, while the chapters presenting innovative electronic tools will be discussed in two separate paragraphs, using the treatment of linguistic data within these tools as a basis of comparison for the different lexicographical practices.

## 2. Fuertes-Olivera and Bergenholtz's *e-Lexicography*

The book by Fuertes-Olivera and Bergenholtz presents fifteen contributions ranging from theoretical issues in lexicography to the illustration of innovative dictionary projects and the features that dictionaries will have in the future. Particularly significant is the inclusion of two chapters (4 and 5), that have Information Science as their topic, since one of the crucial claims of the Aarhus School is that lexicography should be an autonomous discipline within the field of Information Sciences. These chapters follow immediately after the introductory section, dedicated to the theoretical stances of the lexicographical function theory, and they precede the presentation of different e-dictionaries. The *Base lexicale du français*, a particularly promising e-tool for computer assisted language learning, is described in the closing section (chapter 13) along with a usability study (chapter 14) and a brief conclusive chapter (15), containing the main stances discussed at the congress held in Valladolid. Chapter 6 serves as a bridge between theory and practice, and presents "four lexicographic information tools"; while chapters 10 and 11 are concerned with language dictionaries: the first reviewing the online resources for the English language, the other presenting the issues posed by dictionary definitions to the function theory.

After the co-authored introduction by the editors, the book opens with a critical discussion of function theory within the stances of theoretical lexicography. The chapter by Gouws hails the radical shift in the lexicographical paradigm now having the users and the situations of use at its centre. The analysis of the different kinds of functions a dictionary may have with respect to the different contexts of consultation by each single user, allows the lexicographer to identify at least one function that is detached from any language problem. If someone accesses a reference tool in order to acquire new knowledge about something, he is faced with a cognitive problem that has nothing to do with communication, and thus language problems (Gouws 2011: 23). In this respect, lexicography proves to have its own theoretical space, and must not be credited as a linguistic discipline, but as a broader reference science, dealing with dictionaries among many other reference tools, e.g. tourist guides, informative web portals, instruction manuals and so on. After having sketched the theoretical background of the function theory, Gouws underlines that the new reference science demands a radical rethinking, and even "unlearning" of the old practices, dominated by a general "linguistic colonialism". Thus, within this framework, a new dictionary model is proposed, namely that of databanks from which different dictionaries can be extracted when accessing the resource

with specific query systems.

In the second chapter, Bergenholtz complains that the current lexicographical practice neglects to conform properly to the users' consultation needs. What is generally displayed by dictionaries is more suited for language experts than for real users. Access routes are only seldom implemented and there is unsatisfactory literature on dictionary use, since, first of all, it surveys unrepresentative population samples. For Bergenholtz, it is more promising to investigate the access paths of single subjects in task-based testing sessions. He reports on two studies, one of these proves that the fastest and most successful consultations were allowed by those dictionaries that require fewer search steps, something that is achievable only with a proper dictionary customization. The evaluative parameters used by Bergenholtz, such as the time and the goal achievement, are called 'efficiency' and 'efficacy' in Information Science; they are also used by Heid in a study presented in the same book (chapter 14, see below).

Bergenholtz goes on to illustrate the innovative features of the fastest dictionary in the survey, the dictionary called *Meaning of fixed expressions*, which is extracted from a database containing the necessary data to display three other 'monofunctional' dictionaries: one to assist with the use of fixed expressions, one to know more about them, and the other to find fixed expressions with a certain meaning. The 'monofunctional' dictionaries are the innovative tools compiled at Centlex of the Aarhus School: they are tailored to display only what is needed by the user in the access situation and they are extracted by a single database using one of the available search options.

This new dictionary model is indicated by Tarp, in the third chapter, as the only kind of reference work that can be reasonably considered as a product of e-lexicography, since it has been planned from the beginning as an adaptable instrument that users can access and customize to their specific consultation needs. Tarp (2011: 59) calls these dictionaries "T Ford", referring to the well-known anecdote of Henry Ford presenting his new car. Ford complained that if "he had consulted people before inventing" it, they would rather have asked for "faster horses". The claim by Tarp is for a radical shift in the lexicographical paradigm in order to make the best of the electronic medium and present-day technology. However, using these monofunctional dictionaries, a complete customization is not yet achievable. This could only become reality if the data display is completely dynamic, and adaptable to every single consultation need of all the possible users in all the possible consultation situations. If it existed, Tarp would call this dictionary a Rolls Royce.

However ambitious it may be, "modern information technologies can start addressing these issues", as Bothma (2011: 79) declares in the fourth chapter, which is dedicated to the analysis of the challenges of function theory from the point of view of the information scientist. Reviewing the tenets of the theory, Tarp (2011: 69) describes his vision of "the best dictionary in terms of needs satisfaction", namely "any dictionary [...] that allows *either* monofunctional

access *or* individualized access in the framework of its specific and foreseen functions". Therefore he illustrates the current technology that could make this wished-for tool become true: from search and navigation options to filters for user profiling and data selection. In particular, very promising for a fine-grained presentation of contents, are metadata, or markups, which are also used to describe the contents within a document. With such descriptors dictionary customization could improve tremendously, displaying only what is suited to the user; for example, it would be possible to show researchers highly technical entries, and laypeople short descriptions of the same topic extracted from the same databank.

With the aid of all the technologies described, Bothma is even able to figure out the details of Tarp's 'Rolls Royces'. These are tools that users can set according to their preferred profile, and change them on the basis of the situation of use, or according to the desired level of complexity. In this way, the system could automatically improve its adaptation to the user, with reference to his behaviour, and present him only with the necessary information. Moreover, the user would receive recommendations from the system, and could make public or private notes on it.

The next chapter (number 5) by Spohr is as enlightening as Bothma's, since it shows how the theory can be translated into reality: "in the e-lexicography environment, theory informs praxis and praxis informs theory – the one cannot exist or advance without the other" (Bothma 2011: 101).

Spohr presents the software architecture of a sophisticated linguistic database developed using the "so-called *Semantic Web* formalisms" instead of a simple relational database, as there were too many hierarchies and relationships to be encompassed. These pages merit special attention, since first of all the author brilliantly explains the details of software architecture in simple terms to computational novices. Secondly the topic itself is paramount, since it is concerned with the necessary treatment of data in order to achieve the most fine-grained customization of the dictionary contents.

The stored data are linguistic and, in order to elaborate a coherent "structured data model" for them, the component parts of this database have been identified in its "lexical entities" and in the mutual relationships that may exist between them.

The *Semantic Web* tools allow the lexicographer to represent hierarchical structures avoiding redundancy, thus it is possible to extract different information for the same item:

For example, one could say that statements like '*investigation* is a nominalization of *to investigate*' and '*investigative* is an adjectivization of *to investigate*' are more fine-grained ways of expressing that '*investigation* and *investigative* are derivations of *to investigate*'. (Spohr 2011: 107)

The first kind of information ('*investigation* is a nominalization of *to investigate*') could be useful for experts, while the second ('*investigation* and *investigative* are

derivations of *to investigate*) is less technical and could be displayed to lay-people. In the database, the relationship 'is a nominalization of' and 'is adjectivization of' are sub-relations of 'is a derivation of', and the system is able to infer the more general statement about the derivation of the terms automatically. This means much in terms of the customization options that will be available for the lexicographer, and these options can be implemented only by incorporating an "access and presentation model" into the system, namely a "layer" that

defines which of the entities (both classes, properties and instances) are relevant to which users in which situations. [...] In fact, this layer contains not only one but several such filters, each of which may let through or filter out different pieces of information, depending on the specific user and situation types (Spohr 2011: 114).

As a conclusion to this section, the chapter by Leroyer (number 6) deals specifically with the shift in paradigm advocated by the many points of view already presented. After a brief theoretical introduction, Leroyer exemplifies the outcomes of the function theory presenting four different "lexicographically designed information tools": a cancer dictionary (*Lexonco*) aimed at patients and their families customized with three different access modes; a guide for Danish purchasers of French real estate who speak some French and for those having no command of the language whatsoever; a mobile tourist guide with automated and user-driven options; a database that will be improved in order to become an aid instrument for scientific text production. This last project, called ARTES (*Aide à la Rédaction de TExtesScientifiques*) was developed in collaboration with the Université Paris Diderot and is the topic of the tenth chapter in the volume by Granger and Paquot (2012).

From chapter 7 to 10, different dictionary projects are presented that have been developed within the framework of function theory; they will be discussed in section 4 of this review, whereas chapter 8, by Fuertes-Olivera and Niño-Amo deals with some proposals to improve *El Diccionario Inglés-Español de Contabilidad*, which was compiled following the model of the *Accounting Dictionaries* project developed by Centlex of the Aarhus School (presented in chapter 7 by Nielsen and Almind). Fuertes-Olivera and Niño-Amo discuss how to support users more effectively when they are faced with cognitive problems and need to improve their knowledge on a new topic, such as accounting. Their idea is to use hyperlinks to external contents and to provide the dictionary with a "systematic introduction".

In chapter 11, Lew (2011) writes an overview of the current online lexicography for the English language. He addresses "general English dictionaries", "learner's dictionaries", "user-involved dictionaries", "diachronic (historical) dictionaries", "subject-field dictionaries", "dictionaries with restricted macrostructures", "dictionaries with restricted microstructures" and "onomasiological dictionaries". Moreover, he discusses general issues of online lexicography,

such as the so-called "dictionary aggregators", the "step-wise access" facility (Hulstijn and Atkins 1998), the customization options currently available in electronic dictionaries, and the multimedia facilities with which they are provided – graphics, audio and video files. The chapter ends with a report on the resources that have a "dictionary-like interface" (such as *ForBetterEnglish.com*, *JustTheWord* or the DANTE project) which is generated by sophisticated wrappers that conceal the data sources (databases or corpora) and are perceived as true dictionaries by lay web surfers.

Customization proves to be particularly promising also in the field of Computer Assisted Language Learning (CALL), as demonstrated in chapter 13 by Verlinde, who presents the new access paths for the *Base lexicale du français*. This web resource was initially created as a database for the French language but it has been progressively developed into a rich multilanguage assistant which is currently accessible with a device called *Interactive Language Toolbox*, presented in chapter 8 of Granger and Paquot (2012).

In Fuertes-Olivera and Bergenholtz (2011), Verlinde describes three applications of the *Base lexicale du français*, namely a task-oriented assistance with reading, translation and writing. The applications for these tasks require that a text is submitted to the system, thus all the options available are somewhat interactive, managed with pop-up boxes which contain translations or suggestions for more appropriate word choices. Moreover, the text can be analyzed by the system and thus complex lexical items, such as multiword expressions and collocations, can be automatically detected. Another valuable feature of the *Base lexicale du français* is that the underlying database is enriched with a corpus of academic words and one of learners' texts, thus the translation assistant can suggest the correct stylistic variant for academic terms, while the writing tool can identify the syntactic and lexical "problematic patterns" for learners, with special reference to Dutch native speakers whose specific errors have been added.

In the last contribution before the closing summary of the Valladolid Congress (chapter 15), Heid (chapter 14) reports on the results of a "set of experiments on usability testing of electronic online dictionaries" (Heid 2011: 288), which were carried out in laboratory sessions according to the usability tests protocols used in Information Sciences. Small groups of participants are required, since "good practice has shown that 12 to 15 lay testers of a homogeneous user group will be sufficient to identify the majority of usability problems" (Heid 2011: 293), while the analysis is made with reference to all the actions performed by the subjects in task-based test activities. Heid (2011: 298) describes how detailed the protocol is:

The sessions were carried out in a usability laboratory using the Morae Observer software for each task, and each subject, keystrokes, and mouse movements were recorded, to get a picture of the navigation behaviour of the subjects; furthermore, screen video and sound recording were used to capture think-aloud protocols.



During the experiment an observer also takes note of many aspects of the session development.

The aim of these investigations is to assess the "usability" of software products, namely, if they are effective when the users are given what they need, and if they are efficient when the effort to satisfy users' needs is adequate<sup>1</sup>.

Even though these tests gave evidence of the very basic consideration that "simple search interfaces work better than more complicated ones" (Heid 2011: 300), it was found during the laboratory sessions that some access routes of the *BLF* are malfunctioning, while others work perfectly well.

Also promising are Heid's ideas about possible further developments with respect to mock-ups that could allow investigations upon single critical aspects rather than whole consultation processes.

### 3. Granger and Paquot's *Electronic Lexicography*

The first part of the volume by Granger and Paquot deals with "Lexicography at a Watershed". It opens with an overview by Rundell of the current editorial policies in dictionary-making. He explains that the shift from p- to e-dictionaries is still an ongoing process for which a specific "business model has not yet emerged". However, Rundell's point of view is far from nostalgic, since "dictionaries have at least found their ideal platform in the online medium" and the corpus revolution has brought forth incredible advances in language descriptions in terms of the "reliable generalizations" which are allowed about lexical patterns and meaning discrimination. Rundell sketches the main steps of dictionary-making regarding the new tools and technologies available that are definitively going to turn the lexicographer's work from the "origination of dictionary text to the validation of decisions made by the software" (Rundell 2012a: 28): from the software for lexical profiling, to workflow management tools and proformas that provide the outline of the entries.

The volume continues with a chapter (number 2) by Kilgarriff and Kosem, who present the current corpora technologies, with special reference to the Sketch Engine. Their contribution is conceived as a guide to the available resources and facilities provided by these tools. Particularly noteworthy for the development of corpus linguistics are the advantages provided by the 'word sketches' of the Sketch Engine, "one-page automatic, corpus-based summaries of a word's grammatical and collocational behavior" (Kilgarriff and Kosem 2012: 44). Since its first use for the Macmillan Dictionary, it has changed the lexicographer's view of large corpora, because "higher volumes of data helped to make the sketches an ever-more reliable reflection of real usage, but without adding to the lexicographer's workload" (Rundell 2012a: 22).

The overall importance of the corpora revolution on lexicography is stressed and analyzed by Hanks (chapter 4), who speculates about what the dictionaries of the future will look like: "contextualization and phraseology will

come to take centre stage. These dictionaries will be electronic products with hypertext structures and links" (Hanks 2012: 64). Whatever the case may be, the chapter presents Hanks's view of lexical meaning, which underlies all his considerations and criticisms, namely the fact that words have only

'meaning potential' rather than meaning as such, and that, at least in the case of verbs, words need to be put into context before any attempt is made to define their meaning. Different contextual patterns activate different components of a word's 'meaning potential'. (Hanks 2012: 68)

According to Hanks, these findings should also influence the lexicography of the future, whose descriptions will be dependent at least on the word classes. Nouns for example could still be described on the basis of their referents, but as for verbs or adjectives only "their normal phraseology" will count. This implies that lexicography, being based on the analysis of the real use of words, is "in a position to spread [...] radical new approaches to the theoretical understanding of meaning in language" (Hanks 2012: 76). Starting from these premises, the *Wiktionary*'s review by Hanks is sometimes tinted with humour:

Our sympathies may be with an anarcho-syndicalist approach to lexicography, but it is hard to imagine how a radical new approach to defining verbs or natural-kind terms and thus enhancing our understanding of the nature of language and meaning could be carried out systematically by large, uncoordinated groups of enthusiasts and volunteers, some with more expertise than others. (Hanks 2012: 82)

The next chapter (number 5) by Abel covers the management process of dictionary editing, and describes the main components of dictionary writing systems: the editing tool, providing the entry template, the database and "administrative tools", used for the whole project management, from the assignment of duties to the monitoring of the dictionary text. In the last paragraph, Abel reports on the latest uses of databases for the compilation of dictionaries. She mentions cases in which the database is the dictionary itself, such as the DANTE project (as it is also reported by Lew 2011, see above), and, vice versa, of multiple dictionaries extracted from the same database; the reference is of course to the "pluri-monofunctional tools" compiled by the Aarhus School, and in particular to the sophisticated databank system by Spohr (2012, see above).

Tarp follows Abel in chapter 6. He underlines the role that abstraction plays in reconsidering the nature of the needs that lexicography aims at satisfying: apart from those needs which are specific to every single dictionary, the general ones are simply 'information'. Therefore it is necessary to turn from the pure "art and craft" of lexicography to a general theory, in order to derive a comprehensive paradigm to design not only valuable present-day tools, but also those of the future. Abstraction has come into play in function theory also for profiling users' needs, since for a science it is necessary to elaborate valuable generalizations at least on its chief standpoints. Users were thus consid-

ered as 'types' for a while, and their dictionaries were the 'monofunctional' tools set for each of them, but Tarp (2012: 115) expresses his dissatisfaction with this simplistic generalization, since

no *type* of user has ever made a *type* of lexicographical consultation in order to access a *type* of data that may meet a *type* of information need occurring in a *type* of social situation. The only thing that has ever happened, and which happens every day, hour, and minute, is that an *individual* user with *individual information needs* occurring in an *individual situation* decides to make an *individual lexicographical consultation* in order to access the *concrete data* that may satisfy his or her individual needs.

The shift from the 'kind' of user to the real one implies a complete rethinking of dictionaries, from the 'monofunctional' to a more customizable tool, whose features can be changed manually by the user, or set automatically by the system detecting the user's behaviour. This is how function theory is going ahead now with its vision of future electronic reference tools.

Prinsloo (chapter 7), dealing with "Electronic Lexicography for Lesser-resourced Languages" and taking the South African context as its focus, complains that the situation is poor and, though the current technology could do a lot to manage the demanding morphology of the Bantu languages, there are still very limited resources for them.

The second part of the book, dedicated to "Innovative Dictionary Projects", starts with the presentation of the *Interactive Language Toolbox* (ILT) by Verlinde and Peeters (chapter 8), which is the new interface of the *Base lexicale du français*, designed on the basis of a thorough usability study which was also reported on by Heid in the volume by Fuertes-Olivera and Bergenholtz (2011, see above). The need for greater simplicity in order to satisfy users' needs has been converted into a single search box which allows 'incremental searches' (or "step-wise access", as Hulstijn and Atkins call it, see above) and has been improved in terms of user-friendly interface and layout.

The other dictionary projects, covering chapters from 9 to 12 and 14, are the object of specific analyses in the next section, while chapter 13, by Meyer and Gurevych, reviews three language versions of *Wiktionary*, the English, German, and Russian. The survey compares the collaborative Wiki platform to expert-built lexicons, testing terms and senses coverage, definitions, and indications about register and style. The findings of Meyer and Gurevych are very different from the considerations made by Hanks in the same volume (see above), since in these respects, the resource proved to be valuable; moreover, it offers a rich hyperlinking system, different access routes and illustrative graphics, and therefore it is to be expected that it can serve different types of users and needs.

The user's perspective is the topic of the third part "Electronic Dictionaries and their Users", which opens with a review by Dziemianko (chapter 15) on empirical studies about dictionary use. Unfortunately, research comparing

paper and electronic dictionaries proves to be disorienting not only for the different outcomes, but also with reference to the tasks administered, number of participants, and survey type – from eye-tracking to record sheets. In her conclusions, Dziemianko underlines that even for the few tendencies that emerge, such as positive vocabulary retention when there is more search effort, other investigations are needed. Her final claims are in line with those by Heid (2011, see above), stating the need for "prefabricated, purpose-built dictionaries" in order to control the parameters involved and allow more adequate comparisons.

In the following chapter (number 16), Lew (2012) evaluates the effectiveness of access to dictionary data, considering morphological (viz. inflected forms) and graphematic drawbacks: if the user does not know the right spelling or the base form of a word, how could electronic dictionaries help? Many options available in current dictionaries are illustrated and discussed. Moreover, Lew explains the advantages and disadvantages of multimedia files, urging caution on the use of video files for learning purposes.

Nesi (chapter 17), instead, reports on 'alternative e-dictionaries', or dictionaries that are not edited by the major publishing houses. Her review is focused on Eastern lexicography but the results are significant at every latitude since these resources are particularly popular among language learners but their quality is not guaranteed by experts.

In chapter 18, Bower writes a wish-list for the lexicographer, requiring specific information a translator may be interested in, such as general and specialized terms, phraseologisms, and multimedia files, but also notes on frequencies and warnings about inadequate word use.

Fuertes-Olivera, in chapter 19, reviews the features of free online business dictionaries and, comparing them to the inadequate printed bilingual resources, concludes that these dictionaries can support learners since they offer valuable extra-resources, such as videos, links to articles on specific topics, and updated data. Moreover, Fuertes-Olivera argues against the use of the available bilingual dictionaries not only because they lack coverage of the subject matter, but also because they offer the illusion of clear-cut cross-language correspondences.

The concluding chapter (number 20) by Müller-Spitzer, Kopleinig and Töpel reports on different surveys carried out applying the protocols of social sciences: many participants were involved, and accurate statistical analysis of data was performed. The findings of this three-year, "externally-funded" research has shown that online dictionaries are accessed using notebooks and desktop computers, rather than smartphones and small-screen devices, while the "tab view" modality of presenting dictionary information is the one that users prefer. It also proved the need for explicit instructions when innovative dictionary features are introduced, otherwise users do not realize their benefits.

Concluding their report, the authors address the objections made by Bergholtz and Johnsen (2005) on survey methodology, claiming that if the strict

parameters of social sciences protocols are applied, this research can be as valuable as the others.

#### 4. Where have all the linguists gone?

The sections in both volumes devoted to innovative lexicographical projects offer a privileged standpoint of comparison. They will be analyzed here with special reference to the way they approach the treatment of linguistic data, because this topic has proved to be the crucial concern of the two opposing visions of "practical lexicography" and "function theory". Before presenting these electronic projects, preliminary remarks are necessary.

If compared to the volume by Fuertes-Olivera and Bergenholtz, it is immediately clear that among the pages of Granger and Paquot's book harsh criticisms have turned into more moderate positions, even though neither of the parties recognizes that there is room for both. Tarp paraphrases the metaphor of the "linguistic colonialism" (Gouws 2011: 22), affirming more fairly that it is simply wrong to believe that linguistics can "answer all challenges" of present-day lexicography (Tarp 2012: 118). On the other hand Rundell (2012a: 29), explaining why there will still be dictionaries in the future, quotes Tarp (2008: 40) himself and then concludes: "it is [...] clear that the dictionary is morphing from its current incarnation as autonomous 'product' to something more like a 'service', often embedded in other resources". It is not specified if these hybrid devices have anything to do with the "information tools" of function theory, namely "any tool, no matter what we call them, aiming to satisfy the needs users might have" (Fuertes-Olivera and Bergenholtz 2011: 3) in all possible situations.

Nevertheless, the speculation on the merging of the two stances is not a lazy intellectual exercise solicited by the broad-coverage of Granger and Paquot's book that urges the reader to put all the pieces of the current debate into a tidy picture. This game in speculation is indeed solicited by the lexicographical projects presented in the book, which in one way or another are indebted to the function theory, or at least make reference to it, but also deal with the linguistic issues of data selection and description.

Granger herself seems unable to perceive the contradiction. In her introduction to the book, she presents the topic of customization of data in the dictionaries' entries quoting different scholars who have addressed it, and after having started from a representative of the "Anglo-Saxon tradition" (Tarp 2012: 107), namely Sue Atkins, she ends with the Danish lexicographers Bergenholtz and Tarp, explaining that the function theory of lexicography "underlies several chapters in this volume" and that "users' needs have become a central issue not only for practical lexicography but also for lexicographic theory" (Granger 2012: 4).

Similarly, the dictionary projects presented in the book by Granger and Paquot deal with language descriptions but are also indebted in one way or

another to the function theory. In this respect, it is useful to consider the features of the DiCoInfo project (*Dictionnaire fondamental de l'informatique et de l'Internet*), directed by L'Homme at the Université de Montréal, and of LEAD (*Louvain English for Academic Purposes Dictionary*), directed by Granger at the Université catholique de Louvain, ARTES (*Aide à la Rédaction de Textes Scientifiques*), coordinated by Pecman at the Université Paris Diderot, the *Danish Sign Language Dictionary*, created by the Centre for Tegnsprog in Denmark, and the *Transpoetika Dictionary* managed by the Belgrade Centre for Digital Humanities.

The ergonomic customization of DiCoInfo (*Dictionnaire fondamental de l'informatique et de l'Internet*), a lexical database organized according to the principles and formalisms of Mel'čuk's lexical functions, has been carried out by the researchers of the Observatoire de linguistique Sense-Texte, L'Homme and Robichaud, in collaboration with Leroyer of the Aarhus School. The authors explain that

the adaptation of DiCoInfo to user-friendly representations has combined the resources of sophisticated linguistic encoding, a functional approach to specialized lexicography, and innovative computational programming for efficient data access and presentation. In the functional framework of lexicography (Tarp 2008, and this volume), dictionaries – and this can apply to specialized dictionaries – are considered as products designed for specific purposes, i.e. tools designed to meet information needs, and therefore solely defined according to the functions they are meant to fulfil. [...]

As far as the adaptation of DiCoInfo to user needs is concerned, two types of assistance were identified – assistance in connection with text production in L2, and assistance with translation from L2 to L1 – both belonging to the category of communicative user situations. (L'Homme, Robichaud and Leroyer 2012: 224-225)

In the closing statements, the authors enthusiastically declare that their sophisticated linguistic analysis and data structure have benefited a great deal from the quoted "functional principles" used to present them:

We attempted to show that dictionaries that contain rich linguistic information encoded with a formal system can still become user-friendly tools. (L'Homme, Robichaud and Leroyer 2012: 235)

Linguistic data were also carefully considered for compiling the *Louvain English for Academic Purposes Dictionary*, a "writing and learning-aid tool" (Paquot 2012: 171) for academic English, supporting non-native speakers. The writing and learning functions required a preliminary linguistic investigation in order to collect the necessary data, and the key lemmata for scientific discourse were identified on the basis of a specific study which extended the inventory proposed by the existing literature with high-frequency words – such as *namely*, *compare*, *aim* – expressing "rhetorical functions that are particularly prominent in academic discourse (e.g. give examples, express cause and effect, conclude, and express possibility and certainty)" (Paquot 2012: 171). Moreover, the learn-

ing function required a specific corpora analysis of text productions by non-native speakers, while the dictionary makes extensive use of the six different corpora (four of "expert writing" and two of learner writing) it contains, as explained by Paquot (2012: 184):

Examples of collocations and lexical bundles are automatically extracted from discipline-specific corpora. Corpora of learner writing are used to inform contrastive error notes targeting specific L1 learner populations. Customization is also implemented at the mesostructure level. According to users' profiles, lexical entries are linked to relevant concordance lines in discipline-specific corpora. As the corpus-query tool is fully integrated into the *LEAD*, users also have access to these specialized corpora to search words that are not in the dictionary.

Conversely, Kübler and Pecman complain that the lack of a consistent description of some language phenomena prevents the desired development of *ARTES (Aide à la Rédaction de Textes Scientifiques)*, a bilingual resource of scientific texts hosted by the Université Paris Diderot. The database aims to offer assistance for translators, professionals and linguists with the most fine-grained issues of specialized phraseology, such as "semantic prosody" and different types of collocations, i.e. "generic collocations" and domain-specific collocations, called by the authors "semantic preference". The semantic prosody is particularly challenging for translation, since it refers to the connotative aspect of meaning that words acquire because of their collocates, and there are few cross-linguistic correspondences in this respect, as Kübler and Pecman (2012: 203-204) explain:

The English verb *commit* [...] has a negative semantic prosody, as it co-occurs with nouns like *crime*, *murder*, *mistakes*, and *suicide*, all words having a negative connotation. The French equivalent *commettre* also presents a negative semantic prosody, as it also occurs with nouns having a negative connotation, such as *crime*, *délit*, *attentat terroriste*, *vol*, *erreur*, and *faute*. A more thorough analysis, however, shows differences between the French and English sets: the French noun *suicide* cannot be used with *commettre*, as an equivalent of *to commit suicide* (\**commettre (un/o) suicide*), [...] and is best translated by the French synthetic verb *se suicider*. Not all typical collocates of the French *commettre* have a negative connotation: in the expression *commettre un roman/une oeuvre/une pièce de théâtre* [...], the nouns themselves do not have a negative connotation, but through the aura of meaning of the verb *commettre*, they take on a negative connotation, yielding a certain irony.

As for the other phraseologisms considered, Kübler and Pecman distinguish those collocates "which are common to a variety of" specialized domains ("generic collocations") and therefore can be used in all scientific texts – e.g. *my concern here is with, in much the same manner, by contrast* – and domain-specific phraseology, since "evidence that the semantic preference of a lexical item is not the same in different varieties of language is overwhelming". For example, "the verb *run* does not co-occur with the same semantic set in general science and computer science English" (Kübler and Pecman 2012: 203).

Though ARTES offers its users privileged onomasiological access to collocations and provides translational equivalences for them, the database architecture seems to be unable to account for the complex requirements of the semantic prosody, since the system encodes lexical units as wholes and each collocation is stored as a single piece, while it should be the other way around, as the authors implicitly admit: "we can instead decide to encode semantic prosodies and preferences as a specific type of relation between terms" (Kübler and Pecman 2012: 207).

More than a thorough linguistic theory, what appears to be missing is a proper data storage architecture, apparently caused by the ongoing nature of the project, which collects data from Master theses and "was designed to cater for teaching and learning needs", while now it seems to be turning into something else for other user groups. Therefore it is useful to refer to Leroyer (2011), since he is collaborating on the analysis of the lexicographical "functional aspects" of ARTES:

the goal is to design a central database prepared for the import and indexing of several thousand terminological records presently stored in other database systems, and to provide advanced search options as well as flexible editing options. (Leroyer 2011: 138)

Linguistic problems are instead crucial issues in sign language lexicography, and prevent the advances that the electronic medium seems to be offering, as Kristoffersen and Troelsgård (2012: 311) affirm in the presentation of the *Danish Sign Language Dictionary*:

It may be argued that a sufficiently strong base of linguistic research results to provide a dictionary with information beyond a basic description of the form and meaning of the sign vocabulary has not yet been developed.

Though focused on one specific sign language, this paper is also concerned with the general issues at stake for all sign languages, which share more features among them than spoken ones<sup>2</sup>. Kristoffersen and Troelsgård highlight how videos can partially answer the "overriding challenge" of "how to render signs in the absence of a written language", offering monolingual definitions and dictionary entry contents. Instead icons can be used as access routes starting from handshapes and place of articulation (the basic components of signs together with orientation and movement, Kristoffersen and Troelsgård 2012: 302), but also additional parameters can be combined in the search facilities that an electronic dictionary allows. However, more basic linguistic issues prevent an adequate lexicographical description of the Danish Sign Language, such as the treatment of synonymic relationships or even the identification of the base form of a word and the lexical class it belongs to. Since great lexical variation is actually a common feature of sign languages and there is no accepted rule allowing one to identify synonyms among them, the lexicographer needs to decide by himself whether two forms are allophones or different



lexical units with the same meaning. In cases like these, a "consistent" analysis, based on explicit parameters, is needed (see Kristoffersen and Troelsgård 2012: 305). On the other hand, only more research on word classes and base forms of signs can bring forth "firm generalizations" to identify the correct lemma for a dictionary entry.

Kristoffersen and Troelsgård (2012: 302) consider the possible benefits of function theory for the great variety of their target users, and say that

it might be more practicable to consider function theory (Tarp 2008), and describe modern sign language dictionaries not by assigning them to traditional dictionary types, but rather by looking at potential user needs in different communicative situations.

However, for the moment, no user customization is available and the present dictionary is defined as "multifunctional", "comprehensive and user friendly" (Kristoffersen and Troelsgård 2012: 302).

Similarly, the *Transpoetika Dictionary*, a WordNet-based Serbian-English pedagogical vocabulary, will probably be adaptable in the future, thanks to a "layered interface that helps users answer specific questions related to their particular communicative or cognitive needs rather than a unified dictionary entry" (Tasovac 2012: 256). However, for the moment, the project seems to be promising only from a linguistic point of view, since, as the author underlines, the WordNet format offers many advantages for a less-resourced language such as Serbian. First of all, the so-called 'expand approach' allows a systematic description of the language using the cross-linguistic comparison with the English Wordnet, and when lack of correspondence occurs, language-specific synsets can be added. Specific improvements of the original format are instead required by the dictionary pedagogical function, such as grammar descriptions, which can be implemented by additional reference to an external MorphoSyntax database, and labels specifying register, regional forms, taboos and all the linguistic variations of which a learner must be aware.

This brief description of the lexicographical projects presented in Granger and Paquot's book shows how many of them are basically concerned with linguistic issues but also consider the possibility of user-friendly access routes to their data. On the contrary, the electronic dictionary tools presented in the book by Fuertes-Olivera and Bergenholtz (2011) take the linguistic data for granted, with the one exception of Spohr's contribution which, however, is focused on how these data are computationally organized. Consider, for example, Nielsen and Almind (2011) in the chapter discussing the route "From Data to Dictionary" (Fuertes-Olivera and Bergenholtz 2011, chapter 7).

Their standpoint is exemplified describing the *Accounting Dictionaries*, one of the leading projects of the Aarhus School, comprising five different monofunctional dictionaries<sup>3</sup>, whose declared "theoretical basis [...] is not text linguistic" but the "lexicographical functions" (Nielsen and Almind 2011: 141), or the kind of help a dictionary should provide to its users in order to satisfy their

needs.

The focus is therefore on the "technical options for accessing" data, not on the data *per se*, since these must be "presented in such a way that users can easily turn [them] into useful information" (Nielsen and Almind 2011: 155). User accessibility is thus a cognitive matter, since

printed and computerized dictionaries contain data and not information, but lexicographers must collect and present data that dictionary users can readily convert into information. The traditional linguistic and text linguistic approaches to lexicography have serious shortcomings, so dictionaries based on these approaches do not fully satisfy the needs for help and knowledge users have in specific types of situation. One way to address this problem is to re-assess the practical and theoretical foundations of online lexicography in light of the electronic options available to produce well-crafted reference tools. (Nielsen and Almind 2011: 166)

Nielsen and Almind describe the overall structure of these tools, which are made up of three different component parts: a database, a website, and a search engine. On the theoretical side, this structure is presented as a "triadic setup" (Nielsen and Almind 2011: 147) in which an "interface", or the website, replaces the static p-dictionary and displays all the contents that can be extracted from the database, using the search engine as a mediator between the database and the website. The conclusion is that, in the electronic environment, the dictionary is only an interface, and its "macrostructure has been replaced by a 'data presentation structure'" (Nielsen and Almind 2011: 148). The lexicographer's main task is to organize "how data elements relate to each other", and his/her main concern regards the type of database structure which will allow the most efficient data selection that, through the access routes of the search engine, is displayed to users. As for the data, in "lexicographically relevant databases [they] are of the type *text* and therefore rather uninteresting from a computational point of view" (Nielsen and Almind 2011: 142).

Computational issues are key, but are not the only ones by any means, since data architecture and management depend on the users' needs to be satisfied. This aim is achieved by "profiling" the intended users with the "diagnostic checklist", namely the listing of questions necessary to describe the specific needs that must be satisfied. For example: what is the users' native language?, what level of proficiency do they have in the foreign language? and "at what level do they master the special subject field" of the dictionary? (see Nielsen and Almind 2011: 150). The answers to these questions "will show the competences of the target group of the dictionary" and allow the selection of the necessary data that can effectively support users, particularly when they are likely to lack specific competences. In cases like these, the lexicographer can even assume a "proscriptive approach", and make recommendations to the user (Nielsen and Almind 2011: 165), suggesting, for example, the most appropriate spelling variant among the many used for the same word. The recommendation is made on the basis of a corpus data study, from which the lexicographer

can verify the contexts in which every variant occurs, and thus make the most valuable generalization on the proper spelling of the term.

At this point, corpus issues appear also within the function theory, which could be suspected to be a contradictory theme within the tenets of this conceptual framework, since it is concerned with linguistic problems and, as such, probably to be discredited. However, what Nielsen and Almind do in practice is to change the perspective from which the matter is seen. The "entire process of selecting data for the dictionaries" is ruled by the operative criteria of "relevance", namely "the quality of being directly connected with the subject field in question", which allows to "distinguish [...] data that directly support a lexicographical function" and therefore are "useful lexicographical data":

For example, collocations are selected because they are important when producing accounting texts, because they are important when translating accounting texts (and often difficult to translate between the languages concerned), and examples are selected because they specifically show how to write and translate accounting texts as well as provide data for knowledge building. (Nielsen and Almind 2011: 154)

As it appears from the description of this dictionary project, the linguistic data *per se* are not separate from the lexicographer's job, but they are only one aspect of the dictionary editorial process and in no way its starting point. Thus in the paragraph describing how the corpora of the *Accounting Dictionaries* were compiled ("Selecting Data for the Accounting Dictionaries Is a Multi-Stage Process"), the main concern regards the preliminary classifications needed in order to correctly "reflect the structure of the field of accounting" and build the proper corpora from which terms, collocations and all the data are extracted. In order to obtain this faithful representation, a careful analysis was carried out not only on the external limits of the subject field (what must not be included), but also on the internal structure of the subject considered and the corresponding terminological classification; while consulting experts of the field prevented "lacunae in the corpora and lack of factual knowledge" (Nielsen and Almind 2011: 153). In these accounting dictionaries, phraseological units – e.g. their extraction, management or presentation – are not shown as problematic issues, which is, however, the case with the DiCoInfo, the LEAD and the ARTES projects in the three aforementioned chapters in Granger and Paquot's volume. On the contrary, special linguistic attention is paid to homonymy and polysemy, considered to be "important aspects in connection with understanding the meaning of specialized terms". In this passage, the authors speak in strict metalinguistic terms when they address 'word class' as a "syntagmatic criterion":

The syntagmatic criterion 'word class' is used for treating terms as homonyms, so that the dictionaries clearly distinguish between homographs belonging to different word classes, for example the noun *expense* and the verb *expense*. (Nielsen and Almind 2011: 157)

They also go into the details of a purely linguistic analysis when dealing with polysemy:

Morphological criteria are generally used in cases of polysemy, so that homographs that can be both countable and uncountable are treated as being polysemous, such as the noun *authority* ('power to make contracts on behalf of another' (uncountable) and 'governmental agency' (countable), and words of the same word class that have different inflectional paradigms are treated as polysemous. (Nielsen and Almind 2011: 157)

Finally, assuming the standpoint of the "relevance" principle, it is possible to infer that the claim against linguistics is more a stance for a radical shift in the lexicographical paradigm, rather than a crusade against the role that linguistics can have in the correct treatment of language problems.

In this respect, the absence of linguists and linguistic theories during the compilation of the *Danish Music Dictionary*, which Bergenholtz and Bergenholtz underlines in chapter 9 of the same book (Fuertes-Olivera and Bergenholtz 2011), needs closer inspection. In this paper, Bergenholtz and Bergenholtz (2011: 189) declare:

Almost all of those who call themselves lexicographers, however, are of the view that lexicography is a linguistic discipline. This is remarkable. For the music dictionaries that are the subject of this contribution, there was no need for any cooperation with linguists or linguistic theories. Three types of experts participated in the planning and execution:

1. an expert in lexicography
2. an expert in music theory and the history of music
3. an expert in lexicographical databases

Obviously, there are dictionaries for which the co-operation of linguists is required, for example, in a general language text production dictionary [...]. For some dictionaries, but not for all, the co-operation of linguists is required.

The music dictionary is used by Bergenholtz and Bergenholtz to exemplify one of the possible applications of the function theory, namely the editing of 'mono-functional' dictionaries, which are those tailored for a specific task and therefore are expected to accomplish their users' needs better than the general 'poly-functional' dictionaries do. More precisely, the *Danish Music Dictionary* is made up of three tools, one for reception, and one for knowledge, while the other one is classified as a 'polyfunctional' dictionary extracted by wide-ranging search routes in the underlying database. This last one "may even serve as a production dictionary", but, since the amount of information provided in the database seems to be absolutely huge, it appears unlikely that this function could be effectively satisfied. The authors themselves prove to be surprised by some search results, such as the word "darbuka" after having typed the string "Berlioz".

However, resuming the initial statement about the fact that the dictionary

was compiled without linguists or linguistic theories, it seems that the implicit assumption is that they were not used simply because they were not needed, since there are no particular language challenges in this lexicographical project, whose functions are strictly receptive and cognitive. In cases like this, lexicography proves to exist on its own, without linguistics at its side. Moreover, the claims for freedom from linguistics also have political implications: "If lexicography were denied any form of being a science, linguistics would remain a discipline at our universities, but lexicography would not" (Bergenholtz and Bergenholtz 2011: 189), and the academic research in this field could be put into question. The metaphor introduced by Gouws of the "linguistic colonialism" in chapter 1 is also in line with these stances, and alludes to the broadening of the lexicographical horizons implied by the paradigm shift of the function theory. Dictionary compiling, as Gouws (2011: 27) states, is a task to be undertaken by a team, whose "members will typically come from different fields, depending on the type of dictionary to be compiled, and these different team members will acknowledge the diversity residing in and implied by the title of 'lexicographer'". In this vision, "linguists are not the only, or even the major, participants in the discussion".

The editorial project of a phrasal verb tool offers additional evidence of the choices and solutions adopted with respect to language data in the function theory. Andersen and Almind (chapter 10) created a database of English phrasal verbs from which three monofunctional dictionaries can be extracted to support Danish speakers with reception, translation, and production of English texts. There is no mention in this paper of any corpus from which the data were collected, while the authors claim to have referred to the British and American lexicographical practice of including in the lemma inventory not only combinations of verbs and adverbs, which is the strict "linguistic definition of phrasal verbs", but also combinations of verbs and a preposition (*to look into the matter*), or verbs with an adverb and a preposition (*to get on with it*). Andersen and Almind (2011: 217) admit that the data are not their concern, since they are the "traditional ones", "the innovative aspect of the project is that the database will be capable of generating three different dictionaries with three different functions". The database is described in the paper in one paragraph dealing with its structure and with the user interface, whereas other interesting outcomes of function theory on the linguistic data treatment can be found with respect to the decisions made on the "sense ordering, grammatical information and the style labels". Firstly, the lemma inventory was enriched with fixed expressions and free combinations that may be problematic for non-native speakers to understand, translate or use. The same criteria guided the ordering of the phrasal verb meanings within the displayed entries. For text reception and translation, it was assumed that the idiomatic meanings should come first, since these are "more likely to be consulted" (Andersen and Almind 2011: 223); instead, for text production the opposite holds: phrasal verbs with transparent compositional meaning are more commonly used by non-native speakers, who

would probably check their syntactic behaviour or recommendations about style and register. However the authors warn that the best solutions to address users' needs must be verified by specific empirical studies and analysis:

It must be emphasized that these observations are not based on empirical investigations, and any decisions about conscious sense ordering will have to rely on findings from investigations into these matters. (Andersen and Almind 2011: 223)

The topic of frequency is also discussed critically by Bergenholtz and Bergenholtz in chapter 9, with reference to the selection of specialist terms, for which "frequency is not an argument; subject relevance is" (Bergenholtz and Bergenholtz 2011: 193), since even the most rarely used technical terms are necessary in order to give a "systematic insight into" the field, regardless of whether they seldom appear in texts or not. The same is underlined by Nielsen and Almind (2011) for the selection of lemmata in the *Accounting Dictionaries*, in this case experts provided the necessary assistance to avoid omissions.

The challenging topic of tailoring definitions, partially dealt with by Bothma (2011: 90-91) who suggests using metadata markups, is addressed by Sánchez and Cantos in chapter 12. The authors apply the Lexical Constellation Model, a proposal for lexical semantics that they presented in two previous papers. Starting from these researches on corpus data, Sánchez and Cantos (2011: 262) elaborate their theory about lexical units which

result from the clustering of specific semantic features which are perceived as units by the speakers. These units are, however, not isolated entities; they may share part of their features with other lexical units, so that the units intervening in the same set of connections are not fully independent as regards their semantic properties. Such interconnectivity is the very foundation of a lexical constellation.

The "constellation model" seems to be inspired by the tenets of structuralist semantics (see Sánchez and Cantos 2011: 266), however its description would have been worth further explanation and more details, since the authors seem to emphasize the general problems of dictionary definitions and the shortcomings of present dictionaries more than illustrating their own model. Additionally, the example they provide to illustrate their theory is quite misleading, since they explain the word *lion* referring to the scientific taxonomy classification, and explaining that it is "categorized first as an 'animal', later 'a mammal', and then 'a cat'" (Sánchez and Cantos 2011: 262). To these general categories, physical features of the animal are added, e.g. its colour and size, and only in the end they underline what is crucial for the lexicon: "In any case, people may normally never mention that lions are animals or mammals, [...] while less common knowledge will be more decisive for identifying the world around us" (Sánchez and Cantos 2011: 263). However, even if the *lion* "may be taken as an illustrative example of the model [they] have in mind", when they "illustrate

[their] proposal at work", the methodology they use is by no means based upon a combination of taxonomies and pieces of cultural knowledge, but resorts to corpora analysis:

*Mano*, 'hand' in Spanish, is defined in a paper dictionary, GDUEsA, with 17 different meanings and more than 100 idioms and set phrases, in which *mano* appears as the head word. We looked for the meanings of *mano* in the corpus *Cumbre* and then we organized them in accordance with the LCM. (Sánchez and Cantos 2011: 266)

The aspect that seems to fit well into the function theory is the customization options that the semantic features implied by the theory allow, since it is possible to link a bundle of them to a particular user profile and thus to create definitions that can be increased and detailed with respect to the intended users.

## 5. No conclusion available yet

The careful analysis of the innovative dictionary presented in the books has underlined a difference in the agenda of the compilers, rather than a difference in their concerns. While all the projects of Centlex of the Aarhus School were conceived from the beginning in accordance with the parameters of the function theory, others started as lexicological archives and then evolved into something else: consider ARTES, but also the *Base lexicale du français* which has been improved step by step to reach smart personalization options. On the contrary, for those that still have major linguistic concerns, such as the *Danish Sign Language Dictionary*, the customization is only one of the many issues to be considered for future improvements.

However, as Spohr's challenging database project demonstrates, customization depends on data storage organization: the more detailed the analysis and archiving of these data, the better they will be displayed to the final user. When the data are linguistic, the general ontology and relational system must be of a linguistic kind, and linguistic theories will undoubtedly help a great deal.

This does not mean that the dictionary must display frequency labels and lists of collocations to school children, but that with one single database, provided with the finest language description, lexicographers would be able to compile one dictionary for linguists and one for primary school students at the same time. More precisely, these dictionaries will not be compiled, they will be displayed, since they will not exist before the user accesses them.

This dynamicity is one crucial point of function theory, since there was a radical shift in the 'user paradigm': the 'type' of user no longer exists, the future is only about real users (Tarp 2012, see above). The electronic tools that will be produced for them will allow the most tailored searches on the basis of manual settings by the users and automatic detection of the user behaviour by the system.

Within the framework of the function theory, one supposes that also empirical studies will play a key role, as claimed by Andersen and Almind (2011: 223), since not all the choices made by the lexicographer in order to satisfy a specific function are obvious, many of them remain arbitrary (e.g. sense ordering, specific entry contents and so on) and based on good common sense. Generally speaking, dictionary testing is one of the great concerns of researchers (Bergenholtz 2011: 30-53; Dziemianko 2012: 319-342), particularly with respect to the use of lexicographical resources by learners, and since there is a general discontent regarding the amount of research that has been carried out in this field, the proposal by Heid (2011: 287-304) of using 'usability testing' protocols seems to be at least new and worthy of further study.

To conclude, it must be underlined that the principles of function theory are taken into consideration in all but three chapters of Granger and Paquot's book, and they are also used in four innovative dictionary projects out of the six presented in the book, so one may suspect that function theory is the current academic fashion, or that Bergenholtz's adage is valuable: "nothing is more practical than a good theory".

All in all, the two volumes are about the future of lexicography and only time will tell what the outcomes of the current research in the field will be. Until that moment no conclusions on the topic can be written, it can only be observed, as Rundell says, that "we live in interesting times".

## Notes

1. For the proper formulation of these concepts, see Heid 2011: 289-291.
2. This situation has been exemplified by Newport and Supalla (2000: 109) with the *dinner conversation paradox*: "A long dinner among Deaf users of different sign languages will, after a while, permit surprisingly complex interchanges."
3. These are the monofunctional dictionaries listed by Nielsen and Almind (2011: 154): *English Accounting Dictionary*, *Danish Accounting Dictionary*, *English–Danish Accounting Dictionary*, *Danish–English Accounting Dictionary*, and *English–Spanish Accounting Dictionary*.

## References

- Abel, A. 2012. Dictionary Writing Systems and Beyond. Granger, S. and M. Paquet (Eds.). 2012: 83-106.
- Andersen, B. and R. Almind. 2011. The Technical Realization of Three Monofunctional Phrasal Verb Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 208-229.
- Atkins, B.T.S. 1993. Theoretical Lexicography and its Relation to Dictionary-making. *Dictionaries: Journal of The Dictionary Society of North America* 14: 4-43.
- Atkins, B.T.S. and M. Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford/New York: Oxford University Press.
- Béjoint, H. 2010. *The Lexicography of English*. Oxford: Oxford University Press.
- Bergenholtz, H. 2011. Access to and Presentation of Needs-Adapted Data in Monofunctional Inter-



- net Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 30-53.
- Bergenholtz, H. and I. Bergenholtz.** 2011. A Dictionary Is a Tool, a Good Dictionary Is a Monofunctional Tool. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 187-207.
- Bergenholtz, H. and M. Johnsen.** 2005. Log Files as a Tool for Improving Internet Dictionaries. *Hermes. Journal of Linguistics* 34: 117-141.
- Bothma, T.J.D.** 2011. Filtering and Adapting Data and Information in an Online Environment in Response to User Needs. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 71-102.
- Dziemianko, A.** 2012. On the Use(fulness) of Paper and Electronic Dictionaries. Granger, S. and M. Paquot (Eds.). 2012: 319-342.
- Fuertes-Olivera, P.A. and H. Bergenholtz.** 2011. Introduction: The Construction of Internet Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 1-16.
- Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.).** 2011. *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. London/New York: Continuum.
- Fuertes-Olivera, P.A. and M. Niño-Amo.** 2011. Internet Dictionaries for Communicative and Cognitive Functions: *El Diccionario Inglés-Español de Contabilidad*. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 168-186.
- Gouws, R.H.** 2011. Learning, Unlearning and Innovation in the Planning of Electronic Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 17-29.
- Granger, S.** 2012. Introduction: Electronic Lexicography – from Challenge to Opportunity. Granger, S. and M. Paquot (Eds.). 2012: 1-11.
- Granger, S. and M. Paquot (Eds.).** 2012. *Electronic Lexicography*. Oxford: Oxford University Press.
- Hanks, P.** 2012. Corpus Evidence and Electronic Lexicography. Granger, S. and M. Paquot (Eds.). 2012: 57-82.
- Heid, U.** 2011. Electronic Dictionaries as Tools: Toward an Assessment of Usability. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 287-304.
- Hulstijn, J.H. and B.T.S. Atkins.** 1998. Empirical Research on Dictionary Use in Foreign-language Learning: Survey and Discussion. Atkins B.T.S. (Ed.). 1998. *Using Dictionaries. Studies of Dictionary Use by Language Learners and Translators*: 7-19. Tübingen: Max Niemeyer.
- Kilgarriff, A. and I. Kosem.** 2012. Corpus Tools for Lexicographers. Granger, S. and M. Paquot (Eds.). 2012: 31-55.
- Kristoffersen, J.H. and T. Troelsgård.** 2012. The Electronic Lexicographical Treatment of Sign Languages: The Danish Sign Language Dictionary. Granger, S. and M. Paquot (Eds.). 2012: 293-315.
- Kübler, N. and M. Pecman.** 2012. The ARTES Bilingual LSP Dictionary: From Collocation to Higher Order Phraseology. Granger, S. and M. Paquot (Eds.). 2012: 187-209.
- Leroyer, P.** 2011. Change of Paradigm: From Linguistics to Information Science and from Dictionaries to Lexicographic Information Tools. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 121-140.
- Lew, R.** 2011. Online Dictionaries of English. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 230-250.
- Lew, R.** 2012. How Can we Make Electronic Dictionaries More Effective? Granger, S. and M. Paquot (Eds.). 2012: 343-361.
- L'Homme, M. Robichaud, B. and P. Leroyer.** 2012. Encoding Collocations in DiCoInfo: From Formal to User-friendly Representations. Granger, S. and M. Paquot (Eds.). 2012: 211-236.

- Newport, E.L. and T. Supalla.** 2000. Sign Language Research at the Millennium. Emmorey, K. and H. Lane (Eds.). 2000. *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*: 103-114. Mahwah, N.J.: Erlbaum.
- Nielsen, S. and R. Almind.** 2011. From Data to Dictionary. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 141-167.
- Paquot, M.** 2012. The LEAD Dictionary-cum-Writing Aid: An Integrated Dictionary and Corpus Tool. Granger, S. and M. Paquot (Eds.). 2012: 163-185.
- Rundell, M.** 2012a. The Road to Automated Lexicography: An Editor's Viewpoint. Granger, S. and M. Paquot (Eds.). 2012: 15-30.
- Rundell, M.** 2012b. "It Works in Practice but Will it Work in Theory?" The Uneasy Relationship between Lexicography and Matters Theoretical. Fjeld, R. Vatvedt and J.M. Torjusen (Eds.). 2012. *Proceedings of the 15th EURALEX International Congress, 7-11 August 2012*: 47-92. Oslo: University of Oslo, Department of Linguistics and Scandinavian Studies.
- Sánchez, A. and P. Cantos.** 2011. e-Dictionaries in the Information Age: The Lexical Constellation Model (LCM) and the Definitional Construct. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 251-274.
- Spohr, D.** 2011. A Multi-Layer Architecture for 'Pluri-Monofunctional' Dictionaries. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 103-120.
- Tarp, S.** 2008. *Lexicography in the Borderland between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tübingen: Niemeyer.
- Tarp, S.** 2011. Lexicographical and Other e-Tools for Consultation Purposes: Towards the Individualization of Needs Satisfaction. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 54-70.
- Tarp, S.** 2012. Theoretical Challenges in the Transition from Lexicographical *p-works* to *e-tools*. Granger, S. and M. Paquot (Eds.). 2012: 107-118.
- Tasovac, T.** 2012. Potentials and Challenges of WordNet-Based Pedagogical Lexicography: The Transpoetika Dictionary. Granger, S. and M. Paquot (Eds.). 2012: 237-258.
- Verlinde, S.** 2011. Modelling Interactive Reading, Translating and Writing Assistants. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011: 275-286.