L2 Writing Assistants and Context-Aware Dictionaries: New Challenges to Lexicography

Sven Tarp, *Department of Afrikaans and Dutch,*
University of Stellenbosch, South Africa and Centre for Lexicography,
Aarhus University, Denmark (st@asb.dk)

Kasper Fisker, *Ordbogen A/S, Odense, Denmark*
(kfi@ordbogen.com)
and
Peter Sepstrup, *Ordbogen A/S, Odense, Denmark*
(pse@ordbogen.com)

Abstract: Dictionaries are increasingly integrated into other tools designed to assist the reading, writing and translation of texts. *Write Assistant* is a newly developed tool aimed at assisting people writing in a second language. It feeds on big data taken in from corpora and digital dictionaries. The paper discusses the philosophy behind the tool, the techniques applied, its empirical basis and functionality, as well as the extent to which it helps its users. It shows how the tool makes it possible to shorten and even skip some phases in the traditional information-search process and allows its user to maintain the focus on the message to be written without the need to consult external information resources. The paper shows how the underpinning technology gives birth to a new type of dictionary that is context-aware and provides a more personalised user service. But it also indicates that future dictionaries need to be conceptionally adapted to the specific tool in order to optimize the service. All this poses new challenges to lexicography.

Keywords: *WRITE ASSISTANT*, INFORMATION TOOL, INTEGRATED DICTIONARIES, L2 WRITING, INFORMATION-SEARCH PROCESS, EMPIRICAL RESOURCES, CORPORA, LANGUAGE MODEL, CONTEXT-AWARE DICTIONARIES, DICTIONARY CONCEPT

Opsomming: L2-Skryfhulpmiddels en konteks-sensitiewe woordeboeke: Nuwe uitdagings vir die leksikografie. Woordeboeke word toenemend geïntegreer in ander hulpmiddels wat ontwerp is om te help met die lees, skryf en vertaling van tekste. *Write Assistant* is 'n nuut ontwikkelde hulpmiddel wat gerig is daarop om mense wat in 'n tweede taal skryf te assisteer. Dit is gebaseer op groot data wat verkry is uit korpora en digitale woordeboeke. In hierdie artikel word die filosofie agter die hulpmiddel, die tegnieke wat toegepas word, die empiriese basis en die funksionaliteit daarvan bespreek, sowel as die mate waartoe dit die gebruikers daarvan help. Daar word aangetoon hoe die hulpmiddel dit moontlik maak om sommige fases in die funksionele inligtingsoektogproses te verkort en selfs te weg te laat, en hoe dit die gebruiker toelaat om fokus op die boodskap wat geskryf moet word, te behou sonder dat dit nodig is om eksterne inligtingsbronne te raadpleeg. In hierdie artikel word aangedui hoe die onderliggende teg-

nologie lei tot die skep van 'n nuwe soort woordeboek wat konteks-sensitief is en wat 'n meer persoonlike gebruikersdiens verskaf. Dit dui egter ook daarop dat toekomstige woordeboekbehoeftes konseptueel by die spesifieke hulpmiddel aangepas moet word om die diens te optimaliseer. Dit hou alles nuwe uitdagings in vir die leksikografie.

Sleutelwoorde: *WRITE ASSISTANT*, INLIGTINGSHULPMIDDEL, GEÏNTEGREERDE WOORDEBOEKE, L2-SKRYFWERK, INLIGTINGSOEKTOGPROSES, EMPIRIESE HULPBRONNE, KORPORA, TAALMODEL, KONTEKS-SENSITIEWE WOORDEBOEKE, WOORDEBOEKKONSEP

1.   Introduction

> To raise new questions, new possibilities, to regard old questions from a new angle, requires creative imagination and marks real advances in science. (Einstein and Infeld 1938: 92)

Over the years, dictionaries have been designed and compiled to meet a big variety of human needs detected in society. One of their many functions has been to assist users who write texts in a foreign language. As Chon (2009) reports, people engaged in this activity frequently consult a combination of L1–L2, L2 and L2–L1 dictionaries when they experience different types of problem and require information in order to solve them. The tradition of consulting a combination of monolingual and bilingual dictionaries has continued from the printed to the digital world where, surprisingly, only a relatively few dictionaries, like *WorldReference*, so far have been conceived as an integrated package that can meet both monolingual and bilingual user needs in connection with L2 text production. Although this last class of dictionary undoubtedly represents an important step forward for lexicography, the time now seems ripe to take another and even bigger step forward in terms of a more personalized and efficient user service. The technological preconditions for this already exist. It is merely a question of taking full advantage of the available computer and information technologies.

One of the many current challenges is to rethink the whole process which people traditionally follow when they look for information. This process is more or less as described by Bergenholtz et al. (2015):

1.   a person experiences an information need in a specific situation,

2.   becomes aware of the need,

3.   decides to take action and consults an information bearer, e.g. a dictionary,

4.   carries out the consultation,

5.   evaluates the result, i.e. if it is satisfactory,

6.   returns to the situation where the information need originally occurred,

7.   and uses the retrieved information to solve the problem that gave rise to the need.

This process is quite complex and demanding. In a lexicographical perspective, it requires, among other things, that the person experiencing the need is aware of this need, can define its character, has a dictionary at hand and knows how to consult it and find the relevant data, is capable of retrieving the required information from these data, is sufficiently prepared to evaluate the retrieved information, and knows how to apply it. Besides, the process is in any case *time-consuming* and *disturbing* as it takes focus away from the main activity which the person is performing, especially if it is a communicative one like text production in a non-native language where the message to be transmitted is the central issue. Nesi (2015: 584) rightly observes that "people typically consult maps, encyclopaedias and dictionaries while they are doing something else". This being the case, any consultation of an external information resource inevitably represents an interruption of the activity in question. It may be assumed that most users of these resources just want to go back, as quickly as possible, to what they were doing in order to maintain the focus.

As an additional complication, writers looking for assistance in one of the many dictionaries available on the Internet are at risk of encountering unforeseen temptations. Many 21st century users, especially young people, expect lookups in such dictionaries to be free. Publishing houses all over the world are therefore struggling to find a sustainable business model, a challenge that have led some of them to resort to advertising as a survival strategy. This may be good for business but not necessarily for their users. Research into the phenomenon indicates that as much as five percent of all visitors to websites using advertising get tempted by the colourful banner ads and click through in order to get further information; see Robinson et al. (2007). Even those who do not follow this road paved with good intentions may still be distracted by the many temptations. When they finally return to the task they were performing they will probably have lost their focus and maybe even forgotten why they started the consultation in the first place.

Hence, from a user perspective the challenge is to design an information tool that makes it possible to shorten or even skip some of the above-mentioned phases and steps in order to reduce the overall consultation time, maintain the writing flow, and avoid that the person in question loses his concentration. In this respect, the key words are *time*, *flow* and *focus*, to which should be added *quality* of the provided information.

In the following sections, we will look at a tool that intends to take up the gauntlet, namely the *Write Assistant* developed by Ordbogen A/S, a Danish IT company providing language services and online dictionaries, both general and specialized. We will discuss the philosophy behind the tool, the techniques applied, its empirical basis and functionality, the extent to which it actually helps its users, as well as the new challenges posed to lexicography in this connection. But before doing so we will briefly look at the relationship between lexicography and technology as well as the requirements to assist L2 text production.

2.   Lexicography and technology: Recent developments

Historically, there is an intimate relationship between lexicography and technology as shown by Hanks (2010), and Rundell and Kilgarriff (2011), among others. Technology has strongly influenced the design, compilation, presentation, distribution, availability, and usage of dictionaries. New disruptive inventions like the printing and computer technologies have led to genuine paradigm shifts with far-reaching consequences for the discipline. Fuertes-Olivera (2016) has described the current situation within lexicography as a "Cambrian explosion". In spite of the undeniable progress that can be observed, lexicography is in many aspects still in the process of fully adapting to the computer and information technologies put at its disposal. Among the main strategic challenges is to work out a lexicographical response to the increasing societal demand for a more personalised service; see Rundell (2010) and Tarp (2011).

Although lexicography can be viewed as a discipline in its own right, it is traditionally characterized by a big interdisciplinary spirit; see Nielsen (2017) and Tarp (2017). This is even more pronounced today where no serious lexicographical project can be carried out without the combined knowledge of lexicography, information technology, linguistics, and other disciplines relevant to the project. No single expert from any of these fields can produce a high-quality lexicographical product on his own. Whoever takes the initiative has to cross the disciplinary border and include knowledge and skills from the relevant fields. The integrated information tools discussed below would never see the light of the day without the combined efforts of IT experts, lexicographers, corpus linguists, Internet designers, etc. In this respect, the central role of lexicography is determined by its long experience and ability to specify user needs, define the corresponding lexicographical data, and establish the best ways to access these data. This role goes far beyond traditional lexicography and dictionary-making as predicted by Tarp (2009).

One of the most promising developments in recent years is the growing challenge to the traditional stand-alone dictionary, whether printed or digital, and the gradual integration of lexicographical products into other digital information tools designed to assist the reading, writing and translation of texts (see for instance Verlinde et al. 2010, Paquot 2012, Verlinde and Peeters 2012, and Granger and Paquot 2015). Another important development affects the very concept of consultation where an increasing number of lookups are made automatically, a phenomenon defined by Tarp (2008: 123) as "passive searching" in contrast to the "active searching" performed by the users themselves. These kind of "passive" lookups frequently take place without the users even being aware that they are consulting a dictionary. Although there is no reliable statistics, today most dictionary consultations are probably made in the various types of integrated information tool and even automatically.

These two innovations point in the right direction. The integration of dictionaries into other tools makes allowance for reduced "information costs" in

terms of the time spent in the consultation and in the processing and application of the retrieved information; see Nielsen (2008). In addition, the introduction of "passive searching" in dictionaries and other information resources tends to neutralize the well-known problem that "most users get tired of consulting a dictionary whenever they encounter an unfamiliar word" (Verlinde and Peeters 2012: 158). Both innovations represent a step forward towards an improved service that leaves the users of these tools with more time to focus on their primary activity, be it reading, writing or translation.

When it comes to writing aid, the tools designed to provide this service can be conceived with different functionalities depending on the underpinning philosophy. The Dutch–French–English *Interactive Language Toolbox*, for instance, is primarily conceived as a language-learning tool and therefore it "does not correct the submitted text" but "only identifies syntactic and lexical patterns that may contain errors" in order to encourage the user "to reflect critically on his writing" (Verlinde 2011: 282-283).

Against this background, a distinction can be made between *detective*, *corrective*, and *predictive* writing assistants. The former refers above all to various types of spelling and grammar checker, but efforts are also being made to develop tools that can check other linguistic categories such as collocations; see Wanner et al. (2013). The main advantage of detective writing assistants is that they can call the users' attention to problems (and needs) of which they may not be aware. By contrast, corrective writing assistants, for example the one provided by Microsoft, do not only detect possible errors in the already written text but also come up with alternative solutions. Finally, predictive tools intervene directly in the writing process with suggestions, either on how to complete a word when the first letters have been typed, or on which word(s) could be next in the sentence. Many people know such interactive and predictive tools from their smartphones and tablets where they have the potential to speed up a writing process usually performed with only one or two fingers.

3.   Information needs in L2 text production

There is a relatively big body of lexicographical literature dealing with the needs people may experience when writing in a second language as well as the response which dictionaries should provide to these needs (see for instance Rundell 1999, Tarp 2004, Bogaards 2005, Chon 2009, and Lew 2016). This literature contains valuable ideas which will be taken into account in the following reflections. Among the mentioned lexicographers there is, for instance, a general understanding that neither monolingual nor bilingual dictionaries are capable of meeting all the users' needs on their own. Only a combination of dictionaries can achieve this. But what combination? And which lexicographical data should these dictionaries offer in order to serve users engaged in L2 text production? An intent to answer these questions will be made in this section. The discussion will be based on the idea that user needs are first of all

determined by the situation in which they occur and then shaped by the relevant characteristics of the person who experiences them; see Fuertes-Olivera and Tarp (2014: 48-57). The starting point for a determination of these needs is therefore an analysis of the L2 writing situation and the different types of problem a writer may encounter in this situation.

If we exclude the already committed mistakes which the writer is not aware of, the most common problems and corresponding information needs are probably the ones included in the following Top Ten where 'word' is used as a generic term that also includes compounds:

	Class of problem	→	Information needs
1	The writer knows an L2 word but cannot remember it.	→	L2 equivalents provided in an L1–L2 solution
2	The writer does not know the word to be used in L2.	→	L2 equivalents provided in an L1–L2 solution with meaning differentiation
3	The writer does not know a collocation to be used in L2.	→	L2 collocations provided in an L1–L2 solution
4	The writer knows an L2 word but is not sure if it can be used with the concrete meaning.	→	Explications
5	The writer knows an L2 word but is not sure if it can be used in a concrete context.	→	Stylistic, pragmatic, cultural or language-political information
6	The writer knows an L2 word but is not sure how it is spelled.	→	Information about orthography (or autocorrection)
7	The writer knows an L2 word but is not sure how it is inflected.	→	Information about inflected forms
8	The writer knows an L2 word but is not sure how to combine it with other words.	→	Explicit and/or implicit information about syntactic properties.
9	The writer knows an L2 word but is not sure how to construct collocations with this word.	→	Collocations with the L2 word in question
10	The writer knows an L2 word but wants to vary the language and use another word.	→	Synonyms and/or antonyms

Schema 1: Frequent problems and information needs in L2 writing

It is a matter of course that the above Top Ten does not give a full picture of all the complex problems and needs in connection with L2 writing, for example the ones related to phraseology and various sorts of fixed expressions. Yet it represents undoubtedly some of the most relevant problems and needs although they cannot be seen isolated from the concrete pair of languages in question. Gender is, for instance, not relevant in modern English but it certainly is in languages like German, Spanish and Italian. To this can be added other classes of problem which people may experience when writing L2 texts in other languages, especially the ones outside the Indo-European family.

It is worth noting the expression "is not sure" used in relation with various classes of problem (4–9). This expression covers two variants of the same problem, namely 1) that the writer does not know the linguistic item in question, and 2) that the writer knows it but just cannot remember it and therefore needs a *reminder*, a very frequent experience for L2 learners who typically master a much bigger passive than active vocabulary and grammar.

As can be observed, only problems of the three first classes listed above require a bilingual L1–L2 solution whereas the remaining problems require a solution based upon L2, i.e. either a monolingual L2 or a bilingual L2–L1 one. Inexperienced writers and learners at the beginner's level may typically encounter more problems of the first three classes than experienced writers and advanced learners, but all of them may from time to time experience all ten classes of problem. The traditional separation of learners into beginners, intermediate and advanced does therefore not seem to be relevant in this connection although it could be argued that children having problems in L2 writing may demand less complex information than adult writers; see Nomdedeu and Tarp (2017).

4.   *Write Assistant*

In the following, we will discuss *Write Assistant* developed by Ordbogen A/S. This tool is aimed at giving instantaneous, high-tech response to the problems referred to in the previous section. As such, it is not designed as an independent text-processing program, but as an *application* which currently can only be used together with Microsoft's Office Package although nothing prevents it from being adapted to other types of software when needed. This means that the writer using the application simultaneously can benefit from the advantages of the Office Package, including the Spelling and Grammar Checker. *Write Assistant* is therefore not designed mainly with detective and corrective functions, but above all with a combination of *predictive* and *translative* functions as we will see below.

Once downloaded, the application can only be activated when Word is the front-most window. It will then appear as an icon in the centre of a Word document, an icon that has to be moved outside the document in order to get activated. From then on, it will work in all Office programs, including Outlook,

PowerPoint, Explorer, etc. In its current version (June 2017), *Write Assistant* is only available for Danish native speakers writing in English but the technology developed allows the incorporation of any new pair of languages within a few weeks provided the right empirical resources are at hand (see below). It is foreseen that the tool will be available in dozens of languages in the nearby future, and to this should be added a series of versions specifically adapted to the terminology of concrete companies. A description and analysis of the tool — including its philosophy, empirical basis, language model, functionality, usefulness, possibilities as well as limitations — is therefore relevant far beyond the Danish borders.

4.1   Underpinning philosophy and empirical basis

Write Assistant distinguishes itself from machine-translation programs that provide more or less adequate solutions as well as from stand-alone dictionaries and corpora that require the users' active decision to be consulted. The tool is not a "one-stop shop" (Bowker 2012: 381), but rather a *delivery-on-demand service*. It is not "a set of components which customers can mix and match according to their needs" (Rundell 2007: 50), but rather a set of components which the tool handles in order to provide *customized service* to its users according to their likely needs in each concrete case.

Write Assistant does not claim to meet all information needs occurring during the L2 writing process. The application is not designed to deal with *what* knowledge is transmitted during this process but only with *how* it is transmitted. If students, for instance, need information on the close relationship between Johann Wolfgang von Goethe and Alexander von Humboldt in order to write an essay on this subject, they will have to look for it elsewhere. Even so, and as its name suggests, *Write Assistant* is not conceived to provide solutions to the writers' communicative needs, but "only" *assistance*. The users of this product are still expected to play an active and decisive role and take full responsibility for the quality of the final L2 text. As an integrated and interactive application, the main purpose of *Write Assistant* is to offer instantaneous assistance which the writers can ignore or accept without wasting time on consulting external information resources, without too many interruptions in the writing process, and without losing their concentration and focus.

Write Assistant is based on big data taken in from two main empirical resources, namely a big L2 corpus and one or more digital dictionaries, among them at least one L1–L2 and another L2-based dictionary. When the first English edition is published (at the end of 2017), the corpus in question will be the British National Corpus with 100 million words. The reason why the application feeds on an existing corpus instead of using data directly from the Internet has to do with the quality of these data. Although continuously actualized and much bigger than any other corpus, Internet contains too much "noise" in the form of unedited texts and misspelled words. *Write Assistant* needs high-qual-

ity data in order to provide high-quality service, and therefore even existing corpora have to be further "cleaned" and items with less than ten occurrences deleted as a means to avoid as many spelling mistakes as possible.

Like the corpora, the dictionaries to be used by *Write Assistant* as data resources have to be available on a digital platform, preferably in the form of so-called Models T Ford that have been designed from scratch for the electronic media (Tarp 2011: 60). In the first edition to be published, the lookups will be made in the biscopal online Danish–English/English–Danish dictionary already handled by Ordbogen A/S and widely used by its subscribers. As such, it represents a good starting point although the incorporation of a monolingual English dictionary is also considered. However, and as will be discussed below, it cannot be excluded that a future need for optimizing the quality of *Write Assistant* may require the design of new, or partially new, lexicographical products that have their data adapted to the very specific requirements of this application.

In summary, it can be said that, apart from the technology developed as an indispensable precondition for any success, the quality of *Write Assistant* depends above all on the quality of the corpora and the dictionaries used as its empirical basis. At this point, all that is required to adapt the application to any pair of languages is the existence of big data in the form of digital corpora and dictionaries. If this empirical basis is at hand, *Write Assistant* can be made available in any language in a very short span of time.

4.2   Language model

In its current beta version, *Write Assistant* performs lookups in either an English *language model* or the biscopal Danish–English/English–Danish dictionary mentioned above; hence, it does not look up directly in the monolingual English corpus. The language model is the result of an automatic analysis and restructuring of the words contained in the corpus. It has been designed to receive *tuples* of up to four words and give each of them a score. We will exemplify this with the sequence "I was drawn forward with the prospect of employment" taken from Samuel Johnson's *The Plan of an English Dictionary* (1747). This sequence consists of a total of six 4-word tuples:

"I was drawn forward"

"was drawn forward with"

"drawn forward with the"

"forward with the prospect"

"with the prospect of"

"the prospect of employment"

Based on the frequency with which they occur in the corpus, each of these six tuples will be given a score that indicates the probability of the respective string of four words appearing together in the English language. The language model based on these principles has three functions: 1) the completion of words, 2) the prediction of the next word in the sentence, and 3) the prioritization of translation candidates.

The *completion of words* will be illustrated with a concrete example. If the writer, for instance, has typed "The man was married to a beautiful wo", the model will single out "to a beautiful" and "wo". These two strings are called *context* and *prefix*, respectively, where the latter should not be confused with the traditional linguistic term. By means of a sorted list of all the words contained in the language model, *Write Assistant* now queries all words beginning with the prefix "wo". This could be *work*, *woman*, *world*, *word*, etc. These possible completions of "wo" will then be added to the context "to a beautiful" resulting in the following 4-word tuples:

> "to a beautiful work"
>
> "to a beautiful woman"
>
> "to a beautiful world"
>
> "to a beautiful word"
>
> etc.

Each of these tuples will be given a score from the model proportional to the number of occurrences in the corpus. The completions in the ten tuples that gets the highest score will be selected and presented to the user in a prioritized order. In the concrete example discussed here, it will be of no surprise that the most likely completion of "wo" is "woman". However, this should not be understood as the final solution, but just a suggestion, as the writer may want to express something else (e.g. dedicate a poem "to a beautiful world") and thus has to play an active role and take responsibility for the word eventually chosen.

If the language model cannot give a score to the 4-word tuple shown to it because it does not contain any such combination of words, the first of the four words will be cut out and the 3-word result presented to the model, and so forth. Hence, "to a beautiful woman" will be reduced to "a beautiful woman", then to "beautiful woman", and finally just to "woman". Each of these quadrigrams, trigrams, bigrams, and unigrams will be multiplied by a specific weight in order to get a score. It goes without saying that the shorter the tuple the smaller the weight used as multiplier.

The *prediction of the next word* in the sentence may, from a user's point of view, represent a welcome assistance. Yet from the programmer's point of view it is just a special case of the word completion explained above. To show this

we will take the example where the writer has typed "When I worked on the oil platform I had a terrible". The language model will now single out the context "had a terrible" and the prefix " " where the latter consists of an empty string. It will then take all the words from the sorted list mentioned above, add them to the context, and give each of the resulting 4-word tuples a score. Similar to the previous example, the additions in the ten tuples that have the highest score will be selected and presented to the user in a prioritized order. In this particular case, the ten words are *time*, *start*, *experience*, *impact*, *effect*, *accident*, *season*, *fight*, *relationship*, and *and*. One of these words may satisfy the writer's needs in a concrete context. If not, an alternative strategy has to be chosen, for example to type an L1 word. The language model will then ensure that the L2 equivalents offered to the user are furnished in a prioritized order.

The following example illustrates the *prioritization of translation candidates*: A writer has typed "The man was gift". The word *gift* is a Danish polyseme and will automatically be looked up in the Danish–English dictionary that contains a number of translation candidates such as *poison*, *venom*, *married*, *toxins*, etc. These words are now used to replace *gift* with the following result:

"The man was poison"

"The man was venom"

"The man was married"

"The man was toxins"

etc.

As was the case above, each of these 4-word tuples will then receive a score from the language model. Fortunately, "The man was married" is the one with the highest score, which suggests that *married* is the most likely translation of *gift* in this concrete context. Once given a score, the translation candidates will be presented to the user in the most likely order.

Hence, what has been generated is a *context-aware dictionary*, i.e. a completely new type of dictionary that marks a further step towards the ideal of a more personalized lexicographical product.

4.3　Functionality and examples

When the user opens a new document — or begins a new paragraph or sentence — with *Write Assistant* running, a small box will pop up with the ten most frequent starter words. If the writer wants to use one of these words, all he has to do is to click on it. The first word will then be marked in green (see Figure 1). The writer has now two options, i.e. either to use this word or to scroll down to the right word using the Down Arrow key. When one of the suggestions is marked in green, the normal function of the return key will be deactivated and this key can instead be used to enter the highlighted word

directly into the text, thus skipping one of the steps in the traditional informa-tion search process discussed in Section 1. Once the word has been added, *Write Assistant* will move forward so that the left edge of the prefix aligns with the left edge of the ten suggestions most likely to become the next word, and so forth (see Figure 2). This also takes advantage of the way the eye works.

The	>
In	>
He	>
It	>
This	>
I	>
A	>
On	>
As	>
There	>

Figure 1:   Starter words

If the writer does not opt for any of the suggested words and starts typing another one, a box with the ten most likely completions will pop up as explained in Section 4.2. Figure 2 shows the example with "The man was mar-ried to a beautiful wo" used previously. However, if a third letter is added and the prefix instead is "wom", this would narrow the field of possible candidates and raise the probability of having the required word among the ones with the highest score.

The man was married to a beautiful wo

woman	>
women	>
world	>
work	>
wooden	>
wooded	>
works	>
wonderful	>
words	>
woman's	>

Figure 2:   The tuple "to a beautiful wo" + completions

The application may also suggest a word with which the writer is not completely familiar, or not familiar at all. In this case, the only thing required in order to get additional assistance is to click the arrow right of the suggestion. *Write Assistant* will then perform an automatic lookup in the English–Danish dictionary and present the result in a new window containing a dictionary article that provides meaning explications, among others. This function can be exemplified with the sequence "I want to give a bidrag", where *bidrag* is a Danish polyseme with various English equivalents such as *donation*, *contribution*, *subscription*, etc. (see Figure 3).
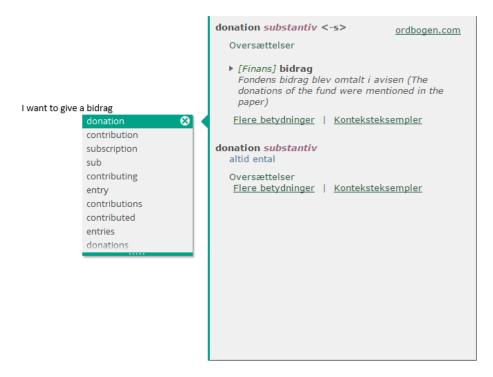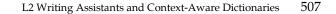

Figure 3:   The tuple "to give a bidrag" + default article

The default article reproduced in Figure 3 furnishes only one sense of *donation*. However, it offers various options to the user who can click on *Flere betydninger* (More senses) or *Konteksteksempler* (Context examples) for additional data (or even on *ordbogen.com* to be referred to the specific article in the online dictionary). If the writer opts for *Flere betydninger*, he will get instantaneous access to more senses as shown in Figure 4.
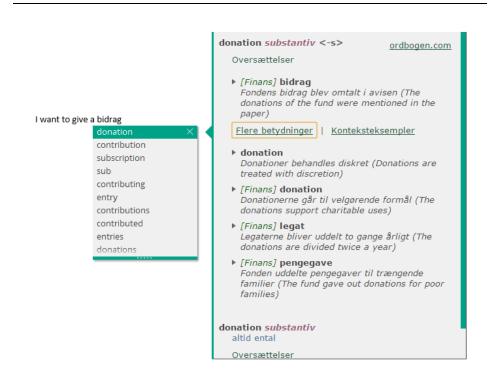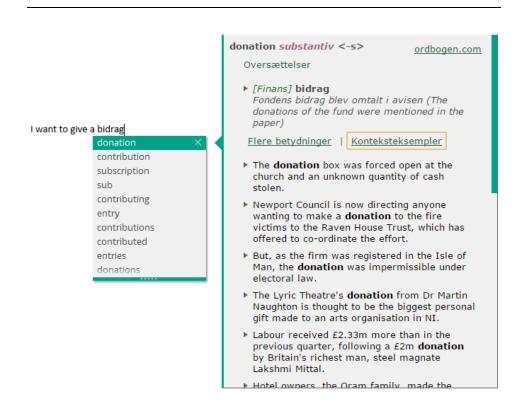
Figure 4:   The tuple "to give a bidrag" + unfolded article

If the problem is that the writer does not know the meaning of the suggested word, or has doubts about it, the information need arising from this uncertainty has all chances of being met immediately. However, the problem may also be that the writer does not know how to use the suggested word together with other words. In this case, further assistance can be achieved by clicking on *Konteksteksempler*. The black data hitherto hidden under this metatext will then be visible in the form of various example sentences uploaded from the English–Danish dictionary and originally selected in a corpus (see Figure 5). This would allow the writer to detect a number of useful collocations, phrases, and syntactic properties, among others, although these data are only provided implicitly. It is worth noting that the metatexts *Flere betydninger* and *Konteksteksempler* are offered with hidden data as default, and that these data can only be unfolded through the user's decision to act. In this way, the increasingly problematic phenomenon of information and data overload, i.e. the "wall of text" that instantly turns people away from webpages, is reduced considerably; see Gouws and Tarp (2018).
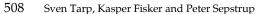
Figure 5:   The tuple "to give a bidrag" + example sentences

Depending on their L2 proficiency level, people writing texts in this language will more or less often have to make use of their mother tongue in order to find or remember the L2 words to be used. Let us take the example where a Danish native speaker wants to say something nice about a woman and does not know, or is not sure of, the English word that can be employed to express this idea. He therefore types the following sequence: "She is very smuk". *Smuk* is a polysemous Danish word with a relatively big number of English equivalents, of which various are more or less synonymous. *Write Assistant* will immediately look up in the Danish–English dictionary and present ten of these equivalents to the writer in a specific, prioritized order (see Figure 6).

She is very smuk
beautiful	>
pretty	>
nice	>
attractive	>
fine	>
fair	>
handsome	>
finest	>
delectable	>
fines	>

Figure 6:   The tuple "She is very smuk" + equivalents

However, if the writer instead wanted to express something similar about a man, the words typed down could be: "He is very smuk". The application would then once more look up in the dictionary and suggest the ten most likely translations of *smuk* (see Figure 7).

He is very smuk
nice	>
handsome	>
attractive	>
fair	>
beautiful	>
fine	>
pretty	>
finest	>
delectable	>
fines	>

Figure 7:   The tuple "He is very smuk" + equivalents

A comparison between Figure 6 and 7 provides another example of the context-awareness of the underlying dictionary. Although identical in this case, the ten equivalents suggested are not presented to the user in the same order. Some of these translation candidates are very close synonyms and the writer's use of one or another would not make a big difference. However, if we look at *beautiful* and *pretty*, these two words are presented as the most likely equivalents of *smuk* in relation to *she* whereas they are relegated to the fifth and seventh position in relation to *he*. By contrast, *nice* and *handsome* are the two highest ranked candidates in relation to *he*, whereas they are number three and seven in con-

nection with *she*. This difference is undoubtedly relevant inasmuch as it reflects real language use as it has been collected in the corpus. In any case, if the writer is still not sure which word to use, all that he has to do is to click the arrow right of the suggestion in order to consult the explications of the respective words and, if it is deemed necessary, click through to the example sentences available under *Konteksteksempler*.

From a philosophical point of view, it is interesting to observe how the relationship between man and machine changes when the user decides to click through to a dictionary article or click on one of the metatexts. Until then, the writer has only been typing letters and words whereas the tool has suggested possible solutions which the writer can ignore or accept with a simple movement of the fingers. Nothing else has been required from the writer. Although interacting with the tool, its user has basically played a passive role in terms of the assistance provided. By contrast, when the user clicks the arrow right of the suggestion, or click on the metatexts, then he turns into an active player that takes a conscious decision to look for further information. This situation is very much similar to the consultation of traditional stand-alone dictionaries, but with one important difference. Whereas the latter have to be taken down from the shelve or accessed on a separate website, the dictionaries integrated into *Write Assistant* are already there with an invitation to be consulted.

4.4   Completely or partially solved information needs

In this section, we will try to answer the probably most essential question, namely to which extent *Write Assistant* actually meets its users' information needs when they experience problems in relation to text production in a foreign language. As a reference we will take the ten typical information needs listed in Schema 1.

(1) When writers think they know an L2 word but cannot remember it, they will need an L1–L2 solution providing *L2 translations* of L1 words. In the previous section (Figure 6 and 7), we saw how the application fully meets this requirement, at least for simplex words. But what about compound words? In an assistant for Danes writing English texts, this word class does not constitute a big problem. Most Danish compounds are written as a single word, and as long as they are selected as lemmata in the L1–L2 dictionary, their English equivalents will automatically be presented to the users when required. However, as most English compounds consist of two or more single words with space in between, a slightly different technical solution would be needed if the tool was designed to assist English native speakers writing Danish.

(2) When writers do not know the English word to be used, they will need an L1–L2 solution providing L2 translations as well as *meaning differentiation*. This requirement is also fully met by *Write Assistant*, but in a way that differs from the traditional solution in dictionaries where meaning differentiators, as a rule,
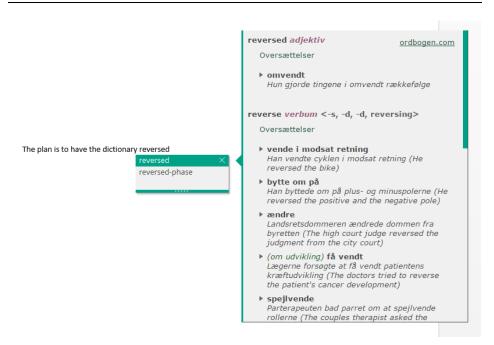
are visible simultaneously. In this application the ten most likely equivalents are furnished in a prioritized order (a big advantage in comparison to traditional dictionaries), but meaning differentiation has to be accessed for each equivalent separately. In this case, the *explications* of the respective L2 words (now presented as equivalents) serve as meaning differentiators (see Figure 4).

(3) When writers do not know a collocation to be used in L2, they will need an L1–L2 solution furnishing *L2 translations* of L1 collocations. In its current version, *Write Assistant* does not offer such a solution. It works relatively well when the collocations are straightforward with similar bases and collocates according to Hausmann's (1985) 2-word collocation theory. If the Danish writer, for instance, types the verb *børste*, the application will suggest *brush* as the first translation candidate, and if the writer then continues with *mine tænder* (my teeth) there would be no big problem. But if L2 was Spanish where the equivalent collocation is *lavarme los dientes* (wash my teeth), there would be no direct solution. But even if there is a more or less similar L1 collocate this would not necessarily mean that the solution was easy. Let us take the Danish collocation *blande kort* (shuffle the cards) as an example. When the writer types the highly polysemous word *blande*, e.g. in the sequence "John was the next to blande", a number of possible English translations of *blande* will be proposed by the tool. However, *shuffle* will only pop up as the fifth most likely candidate after *mix*, *blend*, *mixed* and *dilute* (see Figure 8). The risk of choosing the wrong English equivalent would probably be high if the writer is not very conscious about the problem and looks for additional assistance in the integrated dictionary where *shuffle* is indicated as the right verb in connection with *cards*.

Figure 8: The tuple "the next to blande" + equivalents

(4) When writers know an L2 word but are not sure whether it can be used with the concrete meaning, they will need *explications* of the L2 words. This requirement is fully met as can be seen in Figure 9.
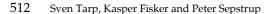
reversed *adjektiv*   ordbogen.com

Oversættelser

▸ **omvendt**
Hun gjorde tingene i omvendt rækkefølge

reverse *verbum* <-s, -d, -d, reversing>

Oversættelser

The plan is to have the dictionary reversed

| reversed | ✕ |

reversed-phase

▸ **vende i modsat retning**
Han vendte cyklen i modsat retning (He reversed the bike)

▸ **bytte om på**
Han byttede om på plus- og minuspolerne (He reversed the positive and the negative pole)

▸ **ændre**
Landsretsdommeren ændrede dommen fra byretten (The high court judge reversed the judgment from the city court)

▸ *(om udvikling)* **få vendt**
Lægerne forsøgte at få vendt patientens kræftudvikling (The doctors tried to reverse the patient's cancer development)

▸ **spejlvende**
Parterapeuten bad parret om at spejlvende rollerne (The couples therapist asked the

Figure 9:   The tuple "have the dictionary reversed" + article

(5) When writers know an L2 word but are not sure whether it can be used in a concrete context, they will need *stylistic, pragmatic, cultural* or *language-political* information. This requirement is only met to a limited degree. The L2–L1 dictionary integrated into the tool does not offer explicit data of this type. But the information may, to a certain extent, be deduced from the example sentences available under *Konteksteksempler*. As a special case, *Write Assistant* will also be available in versions specifically adapted to concrete companies that have their own language policies in terms of the terminology to be used by their employees. These special versions of the application will therefore be designed to reflect and transmit the relevant language-political information.

(6) When writers know an L2 word but are not sure how it is spelled, they will need information about *orthography* or *autocorrection* when typing L2 words. This requirement is fully met. On the one hand, the tool offers correctly spelled completions once the first letters are typed; and on the other hand, it runs together with Microsoft's Grammar and Spelling Checker that will detect possible spelling mistakes and suggest corrections. Alternatively, the user could simply write an L1 word in order to get the rightly spelled L2 word.

(7) When writers know an L2 word but are not sure how it is inflected, they will need information about *inflected forms*. This requirement is fully met. The possible completions, next-words, and equivalents suggested by the application are all inflected in order to fit into the sentence, but only by statistics in the

language model, i.e. without applying grammatical rules. If additional information is required, the full inflectional pattern of the word in question can be accessed in the L2–L1 dictionary (see Figure 9). In any case, Microsoft's Grammar and Spelling Checker will also be there as a more or less effective grammatical safety net.

(8) When writers know an L2 word but are not sure how to combine it with other words, they will need explicit and/or implicit information about *syntactic properties*. This is one of the most complex issues in lexicography. Like almost all existing dictionaries *Write Assistant* does not come up with a completely satisfactory response but it provides at least some assistance. As can be seen in Figure 10, the L2–L1 dictionary offers some syntactic "rules" that are subsequently exemplified. Additional assistance can also be found in the example sentences (in Figure 5), but the users have to deduce the underpinning rules themselves when they want to construct their own sentences. However, neither the explicit nor the implicit data are addressed to specific senses, a weakness that leaves room for mistakes. What is required to satisfy the needs of different user types is a combination of explicit data (rules) and implicit data (example sentences), all of it addressed to specific senses of the word. This requirement is only partially met by the application in its current version.
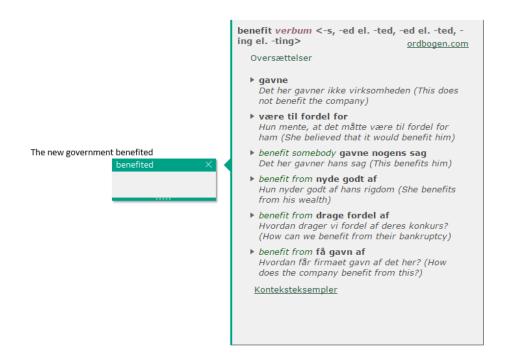

Figure 10: The tuple "The new government benefited" + article

(9) When writers know an L2 word but are not sure how to construct colloca-
tions with this word, they will need *collocations* with the L2 word in question.
This requirement overlaps with the one discussed under Point 3. The main dif-
ference is that the writers in this case know at least one of the words composing
the collection. The problem could therefore be solved by consulting an L2-
based dictionary that contains collocations. As can be seen in Figure 11, the
dictionary integrated into the tool does offer some explicit collocations apart
from the ones that can be deduced from the example sentences. Besides, if the
writers know the first word in the collocation and type this word, then it is
possible that the second word in the collocation will be among those suggested
by the tool, at least if it is a frequently used collocation. If this is not so, or if it is
only the second word that is known to the writer, then an L1–L2 solution is
required as in the example with *shuffle the cards* discussed above. Such a solu-
tion is currently not provided either. Hence, *Write Assistant* has not yet come
up with a completely satisfactory response to the challenges posed by colloca-
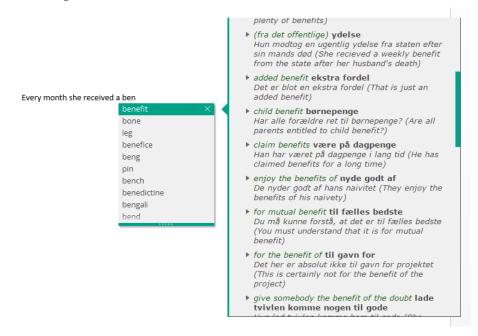tion writing in L2.

Figure 11: The tuple "she received a ben" + article

(10) When writers know an L2 word but wants to vary the language and use
another word, they will need *synonyms* and/or *antonyms*. This requirement is
only met to a very modest degree. When the tool suggests the possible next
words to be written in the sentence, there are sometimes synonyms among the
proposed words but this solution can, by definition, never be systematic inas-
much as the words presented to the user by the language model do not depend

on their semantic structures but on the likeliness of the respective tuples to be found in L2. Thus, a more appropriate solution would be that the underlying dictionary offered synonyms and antonyms, a service it could easily render.

In short, it can be concluded that the first version of *Write Assistant* fulfils half of the ten requirements to a tool designed to assist L2 writing, whereas it meets the remaining requirements to different degrees. The limitations can be due to the technology developed, the design of the interface, or the quality of the empirical basis. Among the five requirements that have not found a satisfactory solution, only those related to L2 collocations can be explained by technological limitations. The challenge is therefore to further develop *Write Assistant* to the extent that it does not only complete or predict words based on the previous 3-word context, but also looks backwards and proposes corrections of already typed words in order to offer the correct collocations. Such a solution would also be relevant to multiword compounds and terms, and would, simultaneously, raise the context-awareness of the integrated dictionary.

The remaining problems detected above — i.e. those related to the treatment of synonyms, antonyms, syntax, style, pragmatics, and culture — require a different solution inasmuch as they have to do with the design of the visualized dictionary articles and as well as the content of the underlying dictionary used as empirical resource. These problems are all of a lexicographical nature and will briefly be discussed in Section 5.

4.5   Usefulness and writing flow

In the introduction to this article we referred to the traditional information-search process as it has been described by Bergenholtz et al. (2015). We stressed the need for a tool that could assist L2 writing in such a way that some of the phases and steps in this process could be shortened or even skipped in order to save time and maintain the writing flow without losing the focus on the message to be transmitted. The analysis which we subsequently made of *Write Assistant* indicates that it is a qualified candidate to become a tool with such properties. As can be expected from beta versions of complex digital products, it has been born with some childhood diseases but they all seem to be curable.

The tool appears to be very simple and easy to use from a functional point of view. Its users can work with their normal keyboard and navigate among the suggested words and perform supplementary lookups in the integrated dictionaries by means of either the mouse or the Alt and Arrow keys. Following the same philosophy, the return key has been reprogrammed so it can now be used to enter words directly into the text once they are marked in green. All this creates the technical conditions for an improved writing flow.

Writers may easier become aware of their information needs when they start typing and the application immediately comes up with suggestions for completions and next words. If they are happy with one of these suggestions,

the only action required from them is to mark the word in question and press the return key. The same is the case when they type an L1 word and instantly get a number of L2 candidates. If further information is required about any of the suggested words, a simple finger movement is all it takes to visualize a dictionary article right away. Once the result is deemed to be satisfactory and the return key pressed, the word in question will automatically be added to the text and solve the problem that originally gave rise to the information need.

As can be seen from the presentation above, some of the traditional phases and steps in the information-search process have become less complex while others have been skipped because the need to access external information resources has been reduced considerably. In fact, the users of the application can go a long way without consulting such resources. They can save precious time and concentrate on the message to be written.

By contrast, while *Write Assistant* undoubtedly makes the writers' job easier, it does not take any responsibility away from them. The users of this tool are still expected to critically evaluate the retrieved information, decide which words to use, and take responsibility for the final text. The users, and nobody else, are the sole authors of the texts produced with the assistance of the tool. The latter is only their handy assistant, but not their co-author.

5.   Challenges to lexicography

The existence of computers and huge databases, the programming of the language model, and the functional and user-friendly design make up the necessary technological and technical conditions for an L2 write assistant like the one described here. However, these conditions do not by themselves guarantee the quality of the application. Once they have been created, the quality of the product depends first and foremost on the quantity and quality of the empirical basis from where the data are taken in, that is, the corpus and the dictionaries.

The requirement to the corpus is basically that it should be relatively big, well-composed, up to date, and sufficiently "clean" in order to reduce the risk of spelling mistakes. But what about the dictionaries? Many of the problems detected in Section 4.4 have to do with the underlying dictionaries. A more detailed analysis of the lexicographical data would confirm this tendency. The criticism of the quantity of L1 words with L2 equivalents, the quantity and quality of the latter, the quality of the definitions of the L2 words, the existence or not of other relevant data categories as well as the quality of these — all this is to a large extent a criticism of the dictionaries used as empirical resources.

It is not difficult to find (or compile) dictionaries that contain most of the missing data categories discussed in Section 4.4. With the possible exception

of appropriate cultural notes, the other items — i.e. stylistic and pragmatic labels, synonyms and antonyms, collocations as well as explicit and implicit syntactic data — are already available in many existing dictionaries, at least in a language like English. However, all these items must be adapted to the specific technical and functional requirements of the tool, an adaptation that may have consequences for the overall dictionary concept and the production methods in new dictionary projects. All this poses new challenges to lexicography.

The usage of already existing dictionaries may create some problems that affect the overall quality of the tool. Today, most dictionaries are still conceived to provide assistance to both text production and text reception without the necessary differentiation. The design and presentation of the different lexicographical items used for both purposes may therefore represent a sort of compromise which is not necessarily the most adequate for L2 writing. Besides, there may be other problems in terms of scopus and directionality. The dictionary used to support the *Write Assistant* is *biscopal*, i.e. it consists of a set of bilingual dictionaries in both language directions with the special requirement that all L2 words in the L1–L2 dictionary must have an equivalent entry in the L2–L1 dictionary. The biscopal solution seems to be the most appropriate in connection with L2 writing, and it is also the one recommended for this purpose by most modern lexicographers (see for instance Lew and Adamska-Sałaciak 2015). Even so, it is important that the dictionary is designed from scratch as a *monodirectional* dictionary. The reason for this fundamental conceptual requirement is that every single version of *Write Assistant* has only one user group in terms of mother tongue, namely native speakers of L1, and that any bidirectionality may interfere inconveniently with the lexicographical data presented to this specific user group.

Here we will briefly discuss *meaning items* as just one example of how lexicographical data that have been taken in from a set of not fully adapted bilingual dictionaries may not be the best solution. In traditional bilingual dictionaries there are two main classes of meaning item in terms of their purposes: 1) the one that is used to differentiate between L2 equivalents in the L1–L2 dictionary, and 2) the one that is used to explain the meaning of an L2 word in the L2–L1 dictionary. Different techniques are applied in each case. In L2–L1 dictionaries, where focus is on text reception, the items most frequently used to explain the meaning of L2 words are L1 equivalents, whereas the items used to differentiate between L2 equivalents in L1–L2 dictionaries are typically L1 synonyms and paraphrases.

In Section 4.4, we saw how the same data category played the role as both meaning explanation and meaning differentiator. This double role may create some inconveniences if it has not been foreseen and taken into account in the moment when the concept for the underlying dictionaries was decided. Figure 12 and 13 illustrate the problem.
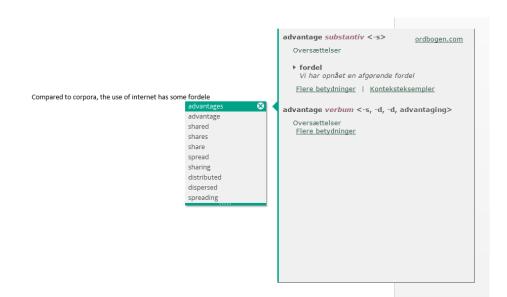
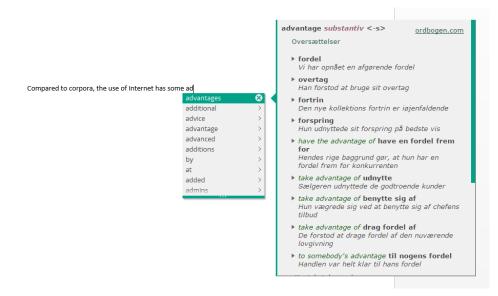**Figure 12:** The tuple "Internet has some fordele" + article

Figure 13: The tuple "Internet has some ad" + article

Although taken from the same underlying dictionary, the two default articles reproduced in Figure 12 and 13 are different because only one meaning item is needed as a *differentiator* in the former, whereas several items are required to

explain the meaning of *advantages* in the latter. In both cases the meaning items consist of an L1 word exemplified by an L1 sentence (e.g. **fordel** — *Vi har opnået en afgørende fordel*). This type of item may be useful when the writers have doubts about the meaning of an L2 word. But they will probably get more confused than enlightened when an L1 word (in this case *fordel*) is used to differentiate between its own equivalents, although the subsequent example sentence may remedy the problem to a certain extent. A different solution is therefore required, preferably a short L1 definition that would be helpful both as an explanation and a differentiator.

Meaning items are no exception. Other classes of lexicographical data also need to be scrutinized and adapted to the special requirements of *Write Assistant* and similar tools as a precondition for optimizing the quality of the service provided.

6.   Perspectives

In the first half of 2017, *Write Assistant* was tested among a small group of Danish upper high school students, and it was also demonstrated to English teachers and students at various Chinese universities. The feedback was generally very positive in terms of its overall usefulness but different opinions were expressed concerning the possible consequences for foreign-language learning. For instance, fear was expressed that future language students may become too dependent on the tool. This may be so. Similar fear was voiced when the calculator was introduced in math teaching. Today it is obvious that many people are highly dependent on this tool, but it is also a fact that it allows them to perform more complex calculations than ever before without committing too many mistakes.

The use of *Write Assistant* and similar tools requires consciousness of the role of man and machine in modern communication. It is still man who is the sole responsible for both the content and form of the message to be written. Technology is only there to assist, not to take over. If this is understood, tools like *Write Assistant* will definitely be helpful in L2 writing. By contrast, the students who currently continue to use machine translation uncritically in spite of repeatedly being warned by their teachers, will probably not be the ones who benefit mostly from these tools.

There is little doubt that the new technology is there to stay. High-tech tools designed to assist the writing, reading and translation of texts will be an integrated part of our lives in the years to come. People will become increasingly dependent on them whether we like it or not. Lexicography can either adapt to this reality or die. Lexicographers are therefore challenged, not only to raise new questions and possibilities, but also to regard old questions from a new angle. There is no perspective in transforming the discipline into a Knight of the Woeful Countenance.

Acknowledgement

References

Bergenholtz, H., T.J.D. Bothma and R.H. Gouws. 2015. Phases and Steps in the Access to Data in Information Tools. *Lexikos* 25: 1-30.

Bogaards, P. 2005. Dictionaries and Productive Tasks in a Foreign Language. *Kernerman Dictionary News* 13: 20-23.

Bowker, L. 2012. Meeting the Needs of Translators in the Age of e-Lexicography: Exploring the Possibilities. Granger, S. and M. Paquot (Eds). 2012. *Electronic Lexicography*: 379-397. Oxford: Oxford University Press.

Chon, Y.V. 2009. The Electronic Dictionary for Writing: A Solution or a Problem? *International Journal of Lexicography* 22(1): 23-54.

Einstein, A. and L. Infeld. 1938. *The Evolution of Physics*. New York: Simon & Schuster.

Fuertes-Olivera, P.A. 2016. A Cambrian Explosion in Lexicography: Some Reflections for Designing and Constructing Specialised Online Dictionaries. *International Journal of Lexicography* 29(2): 226-247.

Fuertes-Olivera, P.A. and S. Tarp. 2014. *Theory and Practice of Specialised Online Dictionaries: Lexicography versus Terminography.* Berlin/Boston: De Gruyter.

Gouws, R.H. and S. Tarp. 2018. Information and Data Overload in Lexicography. *International Journal of Lexicography*. [Advance publication: https://academic.oup.com/ijl/advance-articles.]

Granger, S. and M. Paquot. 2015. Electronic Lexicography Goes Local: Design and Structures of a Needs-driven Online Academic Writing Aid. *Lexicographica* 31(1): 118-141.

Hanks, P. 2010. Lexicography, Printing Technology, and the Spread of Renaissance Culture. Dykstra, A. and T. Schoonheim (Eds.). 2010. *Proceedings of the XIV Euralex International Congress,* Leeuwarden, 6–10 July 2010: 988-1016. Ljouwert: Fryske Akademy.

Hausmann, F.J. 1985. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. Bergenholtz, H. and J. Mugdan (Eds.). 1985. *Lexikographie und Grammatik*: 118-129. Tübingen: Max Niemeyer.

Johnson, S. 1747. *The Plan of an English Dictionary*. https://andromeda.rutgers.edu/~jlynch/ Texts/plan.html. Accessed on 31 May 2017.

Lew, R. 2016. Can a Dictionary Help You Write Better? A User Study of an Active Bilingual Dictionary for Polish Learners of English. *International Journal of Lexicography* 29(3): 353-366.

Lew, R. and A. Adamska-Sałaciak. 2015. A Case for Bilingual Learners' Dictionaries. *ELT Journal* 69(1): 47-57.

Nesi, H. 2015. The Demands of Users and the Publishing World: Printed or Online, Free or Paid for? Durkin, P. (Ed.). 2015. *The Oxford Handbook of Lexicography*: 579-589. Oxford: Oxford University Press.

Nielsen, S. 2008. The Effect of Lexicographical Information Costs on Dictionary Making and Use. *Lexikos* 18: 170-189.

Nielsen, S. 2017. Lexicography and Interdisciplinarity. Fuertes-Olivera, P.A. (Ed.). 2017. *Routledge Handbook of Lexicography*: 93-104. London: Routledge.

Nomdedeu Rull, A. and S. Tarp. 2017. Hacia un Modelo de Diccionario en Línea para Aprendices de Español como LE/L2. *Journal of Spanish Language Teaching*. [To appear.]

Paquot, M. 2012. The LEAD Dictionary-cum-writing Aid: An Integrated Dictionary and Corpus Tool. Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography*: 163-185. Oxford: Oxford University Press.

Robinson, H., A. Wysocka and C. Hand. 2007. Internet Advertising Effectiveness. The Effect of Design on Click-through Rates for Banner Ads. *International Journal of Advertising* 26(4): 527-541.

Rundell, M. 1999. Dictionary Use in Production. *International Journal of Lexicography* 12(1): 35-53.

Rundell, M. 2007. The Dictionary of the Future. Granger, S. (Ed.). 2007. *Optimizing the Role of Language in Technology-enhanced Learning*: 49-51. https://hal.archives-ouvertes.fr/hal-00197203/document/. Accessed on 31 May 2017.

Rundell, M. 2010. What Future for the Learner's Dictionary? Kernerman, I. and P. Bogaards (Eds.). 2010. *English Learners' Dictionaries at the DSNA 2009*: 169-175. Jerusalem: Kdictionaries.

Rundell, M. and A. Kilgarriff. 2011. Automating the Creation of Dictionaries: Where Will It All End? Meunier, F., S. de Cock, G. Gilquin and M. Paquot (Eds.). 2011. *A Taste for Corpora. In Honour of Sylviane Granger*: 257-282. Amsterdam/Philadelphia: John Benjamins.

Tarp, S. 2004. Reflections on Dictionaries Designed to Assist Users with Text Production in a Foreign Language. *Lexikos* 14: 299-325.

Tarp, S. 2008. *Lexicography in the Borderland between Knowledge and Non-knowledge. General Lexicographical Theory with Particular Focus on Learner's Lexicography*. Tübingen: Max Niemeyer.

Tarp, S. 2009. Beyond Lexicography: New Visions and Challenges in the Information Age. Bergenholtz, H., S. Nielsen and S. Tarp (Eds.). 2009. *Lexicography at a Crossroads: Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*: 17-32. Bern: Peter Lang.

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. *e-Lexicography: The Internet, Digital Initiatives and Lexicography*: 54-70. London/New York: Continuum.

Tarp, S. 2017. Lexicography as an Independent Science. Fuertes-Olivera, P.A. (Ed.). 2017. *Routledge Handbook of Lexicography*: 19-33. London: Routledge.

Verlinde, S. 2011. Modelling Interactive Reading, Translation and Writing Assistants. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *e-Lexicography: The Internet, Digital Initiatives and Lexicography*: 275-286. London/New York: Continuum.

Verlinde, S., P. Leroyer and J. Binon. 2010. Search and You Will Find. From Stand-alone Lexicographic Tools to User Driven Task and Problem-oriented Multifunctional Leximats. *International Journal of Lexicography* 23(1): 1-17.

Verlinde, S. and G. Peeters. 2012. Data Access Revisited: The Interactive Language Toolbox. Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography*: 147-162. Oxford: Oxford University Press.

Wanner, L., S. Verlinde and M. Alonso Ramos. 2013. Writing Assistants and Automatic Lexical Error Correction: Word Combinatorics. Kosem, I., J. Kallas, P. Gantar, S. Krek, M. Langemets and M. Tuulik (Eds.). 2013. *Electronic Lexicography in the 21st Century: Thinking Outside the Paper. Proceedings of the eLex 2013 Conference, 17–19 October 2013, Tallinn, Estonia*: 472-487. Ljubljana/Tallinn: Institute for Applied Slovene Studies/ Eesti Keele Instituut.