http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

A Lexicographical Perspective to Intentional and Incidental Learning: Approaching an Old Question from a New Angle

Sven Tarp, *Centre for Lexicographical Studies, Guangdong University of Foreign Studies, China; Department of Afrikaans and Dutch, Stellenbosch University, South Africa; International Centre for Lexicography, Valladolid University, Spain; Centre of Excellence in Language Technology, Ordbogen A/S, Denmark, and Centre for Lexicography, Aarhus University, Denmark* (*st@cc.au.dk*)

Abstract: Second-language learning is a complex process that combines text reception (reading, listening) and text production (writing, talking). Applied linguistics usually distinguishes between intentional and incidental learning. The academic literature contains various definitions of these concepts, especially in connection with reading. The paper explores L2 learning from a lexicographical perspective and redefines the two terms based on parameters like flow, focus, and interruption. It then focuses on digital dictionaries integrated into e-readers, learning apps, and writing assistants, and argues that this integration, so far, has not been particularly successful due to a number of negative factors. As an alternative, the paper provides examples of how lexicographical data could be filtered and presented in pop-up windows to serve both incidental and intentional learning. The former requires instantaneous, contextualized, and discreet assistance with an absolute minimum of lexicographical data, whereas the latter presupposes easy access to relevant additional data. Finally, the paper discusses the techniques and technologies required to guarantee this approach.

Keywords: INCIDENTAL LEARNING, INTENTIONAL LEARNING, INTEGRATED DICTIONARIES, E-READERS, E-READING TOOLS, LEARNING APPS, WRITING ASSISTANTS, INTUITIVE USE, CONTEXT-AWARENESS, LEXICOGRAPHICAL CONTEXTUALIZATION

Opsomming: 'n Leksikografiese perspektief op doelbewuste en toevallige leer: 'n Ou vraagstuk word vanuit 'n nuwe invalshoek benader. Die aanleer van 'n tweede taal is 'n komplekse proses waarin teksresepsie (lees, luister) en teksproduksie (skryf, praat) gekombineer word. Die Toegepaste linguistiek onderskei gewoonlik tussen doelbewuste en toevallige leer. Die akademiese literatuur bevat verskeie definisies van hierdie konsepte, veral met betrekking tot lees. In hierdie artikel word L2-leer vanuit 'n leksikografiese perspektief verken, en die twee terme word, gebaseer op parameters soos vloei, fokus, en onderbreking, geherdefinieer. Dan word daar gefokus op digitale woordeboeke wat in e-lesers, aanleerdertoepassings en skryfhulpmiddels geïntegreer is, en daar word geargumenteer dat hierdie integrasie tot dusver weens 'n aantal negatiewe faktore nie besonder suksesvol was nie. As alternatief verskaf hierdie artikel voor-

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

beelde van hoe leksikografiese data gefilter en in opwipvensters tot voordeel van beide toevallige en doelbewuste leer aangebied kan word. Eersgenoemde vereis onmiddellike, gekontekstualiseerde, en diskrete ondersteuning met 'n absolute minimum leksikografiese data, terwyl laasgenoemde maklike toegang tot relevante addisionele data voorveronderstel. Laastens word tegnieke en tegnologieë wat vereis word om die sukses van hierdie benadering te waarborg, bespreek.

Sleutelwoorde: TOEVALLIGE LEER, DOELBEWUSTE LEER, GEÏNTEGREERDE WOORDEBOEKE, E-LESERS, E-LEESHULPMIDDELS, AANLEERDERSTOEPASSINGS, SKRYFHULPMIDDELS, INTUÏTIEWE GEBRUIK, KONTEKSBEWUSTHEID, LEKSIKOGRAFIESE KONTEKSTUALISERING

1. Introduction

Second-language (L2) learning is a complex process that combines text reception (reading, listening) and text production (writing, talking). In the linguistic and psycho-linguistic tradition, the language-learning process is usually separated into intentional learning and incidental learning; see Krashen (1989), Nagy et al. (1985), Bereiter and Scardamalia (1989), Hulstijn (1989, 2013), Shu, Anderson and Zhang (1995), Laufer and Hulstijn (2001), Brown, Waring and Donkaewbua (2008), and Leow and Zamora (2017), among many others.

The academic literature provides a large number of definitions of these terms, especially in connection with reading. Sometimes, the two terms are treated as synonymous to explicit and implicit learning, respectively, but this approach has been questioned by other authors like Hulstijn (1989: 2633). In any case, the distinctive criterion in the various definitions seems to be the existence or lack of *intent to learn*, frequently in combination with the learner's awareness and consciousness of the process. Hulstijn (1989: 2632) himself defines intentional learning as "a deliberate attempt to commit factual information to memory". Bereiter and Scardamalia (1989: 363) relate it to the "cognitive processes that have learning as a goal rather than an incidental outcome". In this connection, Leow and Zamora (2017: 33) comment that intentional learning "has always been assumed to represent the type of learning, of a more explicit nature, that underscores a formal instructional classroom setting". According to the two authors, the definition of this concept is "relatively stable … albeit with some nuances" whereas there is "quite a range of perceptions" of what incidental learning entails. The different perceptions are "typically reflected in the methodology employed to address its role in the L2 learning process".

Without going too deep into this discussion, a methodological problem in existing research seems to be the lack of terminological distinction between *language knowledge* and *language skills*; see Tarp (2008: 131-136). After reading a text, informants are typically asked what they *know* about a word in terms of meaning, gender, morphosyntax, etc. However, and as Lessing (1747: 8-9) helped us to understand almost three hundred years ago, the purpose of second-language learning is not to obtain some kind of (learned) knowledge of this language,

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

but to develop language skills, that is, "the ability to communicate in the language concerned: to read, write, listen to and speak this language" (Tarp 2008: 132). Learners should not only know words, they should also be able to use them in real life.

From this perspective, another methodological challenge is that many (most?) studies of incidental learning have been conducted in the classroom, or the school context in general, in connection with the informants' reading of written texts. This, of course, provides a more controlled environment to extract reliable empirical data and reach science-based conclusions. But it also excludes other types of social contexts where learners engage in written and oral communication and incidentally pick up words, meanings, and grammatical structures. These contexts also play a relevant role, and sometimes even a crucial one, in second-language learning, especially when the process occurs inside the geographical area where the concerned language is the dominant one. Thus, if the overall learning process has to be grasped in its totality, it seems equally relevant to relate incidental learning to the text production process, especially writing, and to oral communication in general. According to the newest research, modern human beings have existed for about 300,000 years, whereas written language has a much shorter history of 5,000 years for the privileged few and less than two hundred years for the vast majority. This suggests that spoken language has been the primary means of communication for most of their existence. To the extent our ancestors have learned a second or third language, this must have happened incidentally without written texts and formal instruction. Hence, a short excursion into oral communication may be relevant if we want to achieve a broader perspective of the phenomenon of incidental learning.

In the next section, this idea will be used as a background to explore how lexicography can take advantage of current technologies and add a new dimension to the discussion of intentional and incidental L2 learning in relation to digital devices. Section 3 will then look at dictionaries integrated into e-reading tools (e-readers and similar devices) and explore to which degree digital technologies allow us to offer the "ideal solution". Section 4 will follow up with a short discussion of dictionaries that are integrated into learning apps and provide assistance to the reading of L2 texts. The discussion will show how the application of available technologies and techniques already makes allowance for the "ideal solution". Section 5 will deal with multi-word units of meaning and discuss how this challenge can be treated lexicographically in the tools discussed in the two previous paragraphs and, thus, contribute to the two types of learning. Section 6 will move from reading to writing and briefly discuss how digital writing assistants also can contribute to incidental and intentional learning if lexicographical handicraft is combined with cutting-edge technology. Finally, Section 7 will contain the conclusions and sum up how and to which degree lexicography can contribute to incidental and intentional learning.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

2. A lexicographical perspective

In recent years, lexicographers have increasingly emphasised the close relationship between lexicography and information science, and some of them have even categorised the former "as part of" the latter (Wiegand 2013: 14). They have good reasons to do so. The core purpose of all lexicographical products is to assist users with *information* that can meet their needs in different types of situations, among them the ones the Function Theory classifies as communicative situations (e.g. text production and text reception); see Fuertes-Olivera and Tarp (2014: 52).

In the following, we will explore how the provision of information may enhance the learning process in terms of incidental and intentional learning. We will start with incidental learning and illustrate it by means of three examples of oral communication that are well-known to most L2 learners, either inside or outside the classroom, either inside or outside the geographical area where L2 is spoken as a native language.

In the first example (Figure 1), an L2 learner is listening to a native speaker who is describing an experience from the previous day. The learner does not understand one of the words used by the native speaker and asks him to explain its meaning.

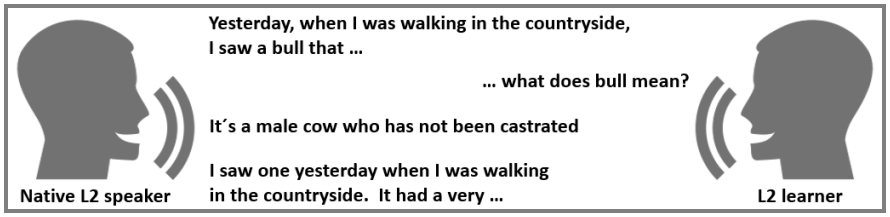

Figure 1: Oral communication between native L2 speaker and L2 learner

In the second example (Figure 2), the roles have shifted. It is now the learner who is describing an experience he had the previous day. Suddenly, he lacks a central word and uses a paraphrase to ask the native speaker to provide the right word.

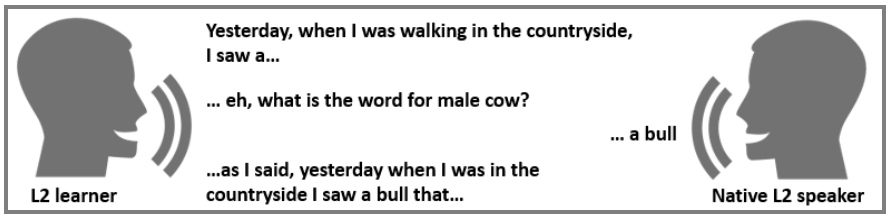

Figure 2: Oral communication between L2 learner and native L2 speaker

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

The third example (Figure 3) shows the learner explaining the same experience, but this time to an L1-speaking teacher of L2. When he does not know or remember a word to express what he wants to say, he therefore addresses the teacher in his mother tongue. The conversation then becomes bilingual and resembles the consultation of L1–L2 dictionaries in connection with L2-text production.

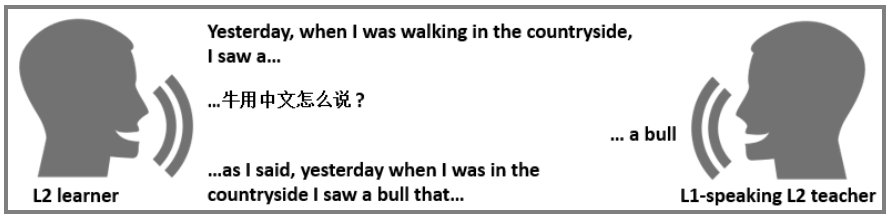

Figure 3: Oral communication between L2 learner and L1-speaking L2 teacher

What do the three figures show? First, they confirm the importance of relating text production and text reception in the language-learning process. Secondly, and apart from the trivial story reproduced, they represent three different types of situations inherent to the topic under discussion. In all of them, the oral conversation between two persons is shortly interrupted. One of the persons does not understand an L2 word in the specific context or does not know which L2 word to use to express his ideas. The other person is then used as a human information resource (human dictionary) that is consulted to get the appropriate information. The short interruption in the conversation fits naturally into the mainline of communication without any of the two persons losing focus on the topic discussed.

The three situations depicted in Figures 1–3 are examples where incidental learning may happen enhanced by the immediate provision of a small piece of information when a communication problem occurs. Although the conversation takes place in the overall framework of the L2-learning process, the learner's *primary intention* in at least two of the three situations is not to learn the second language but to enjoy a normal social conversation.

Inspired by the above examples of oral communication, the concepts of incidental and intentional learning can be adapted to lexicographical consultations performed in connection with the production and reception of written texts. Incidental learning is here related to the situation where learners experience an information need, look up in dictionaries, and get an immediate response that allows them to maintain the reading or writing flow without losing focus on the text and its content and, in this way, pick up new words and meanings. By contrast, intentional learning only starts when the learners interrupt the reading or writing process in order to dedicate time to a deeper study of words, senses, or grammatical structures appearing in the text.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

If the various types of informal chatting are excluded, the writer and the reader are usually separated in time and only interact directly with the dictionary when engaging in communication by means of written texts (see Figure 4). Hence, to achieve incidental learning from the perspective of lexicography, the challenge is to make the consultation process as easy and smooth as possible and reduce the consultation time to an absolute minimum.

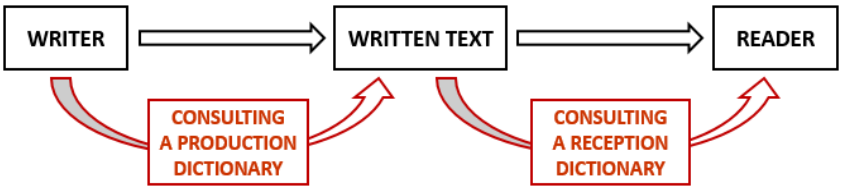

Figure 4: Communication by means of written texts

What happens when learners consult traditional dictionaries, whether printed or app-based ones? First and foremost, they will have to leave the text they are reading or writing and consult an external lexicographical resource. Here, they will frequently get access to a complete article with a large amount of lexicographical data, most of which are totally irrelevant in the concrete context, i.e. information overload. This kind of consultation takes time, disturbs the workflow, and hampers incidental learning. The learners' focus moves from the communication to the consultation.

Hence, the ideal solution seems to be dictionaries that have been integrated into the digital devices learners increasingly use when they read and write. Such dictionaries allow the learners to perform lexicographical consultations without leaving the text they are working with and can be found in e-readers, learning apps, and writing assistants, among others. However, the integration of dictionaries into these devices has not, so far, been particularly successful due to conservative thinking, inadequate adaptation to different types of user needs, imperfect lexicographical databases, poor design of user interfaces, and insufficient application of the available technology. In the following, we will look at both existing integrated dictionaries and those to come.

3. Dictionaries integrated into e-readers and similar devices

We will start with the dictionaries integrated into e-reading tools such as e-readers and similar devices used to read digital texts (e.g. laptops, tablets, and smartphones). These dictionaries are probably the most well-known to users and, at the same time, those that present the biggest challenges in terms of smart technology. They have already been discussed and criticised by many researchers, among them various South African lexicographers, incl. Danie Prinsloo, to

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

whom this article is dedicated; see Bothma and Prinsloo (2013), and Bothma and Gouws (2020).

Compared to traditional dictionaries, the ones integrated into other devices have the advantage that they can be activated by simply touching or clicking on the screen. This technique shortens the initial consultation time and implies that users do not have to leave the book or article they are reading. So far, so good! The problems start when a dictionary article is uploaded to the screen. As an illustration, we have chosen a newspaper article about coronavirus published in Times Live (Singh 2020) and displayed on a laptop. At the end of the article, we have clicked on the word *mark* and activated the integrated dictionary, which is "powered by Oxford Dictionaries". The result is the visualisation of a small excerpt of lexicographical data from the corresponding article (see Figure 5).

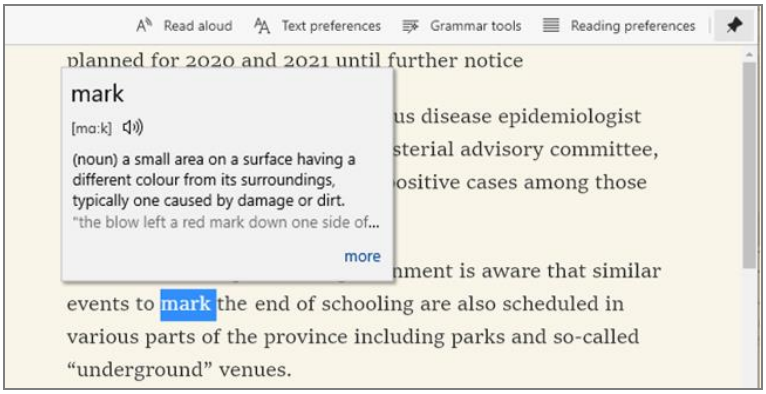

Figure 5: Pop-up window activated on laptop while reading on Internet

It is undoubtedly a good idea that the default pop-up window shown in Figure 5 only furnishes a few lexicographical data. It makes it easier to overview. However, the word *mark* clicked on in the text is a verb, whereas the visualised lemma is a noun. The reader therefore has to make a second click, this time on the signifier "more", to search for the relevant data. The result is a very long article where the user has to scroll down several times. Figure 6 gives an overview of this article after being remodeled. It contains two nouns (based on etymological criteria) and a verb, all of them with two or more senses. The dotted circle highlights the only data relevant in the concrete context. The remaining data are completely superfluous. It will probably take some time to find the required data. As a consequence, the reader's focus will move from the text to the consultation, the reading flow will be interrupted, and incidental learning as defined above becomes impossible. Something has to be done!

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

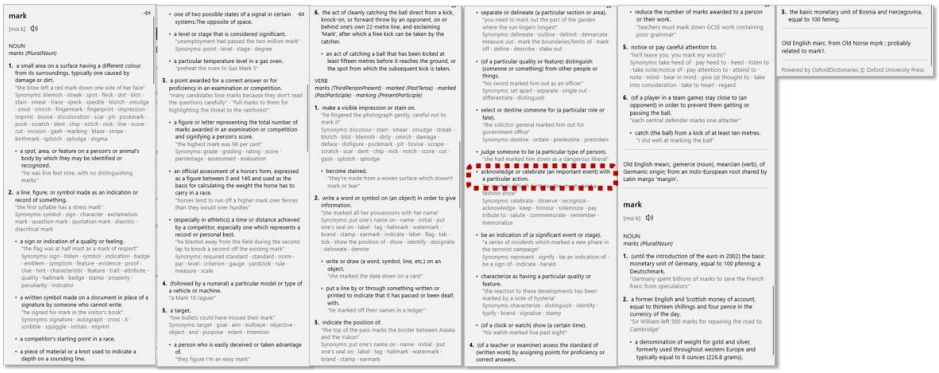

Figure 6: Overview of the complete article after clicking on "more" in Figure 5

A possible solution could be the tagging of word classes, whereby more than half of the lexicographical data shown in Figure 6 could be excluded. The corresponding technology has improved considerably over the past few years, although it is still not completely reliable. As such, tagging could be part of the solution. But even if it were possible to detect the correct word class in the concrete context, the challenge would still be to determine the sense that is relevant in this context (the lexicographical data assigned to the six senses of the verb *mark* represent about 45 percent of the total amount of data in Figure 6). Some of these data may also be superfluous as assistance to reading. But even if they were discarded, there would still be too much data to elegantly fill the default pop-up window activated by touching or clicking on a word in the text.

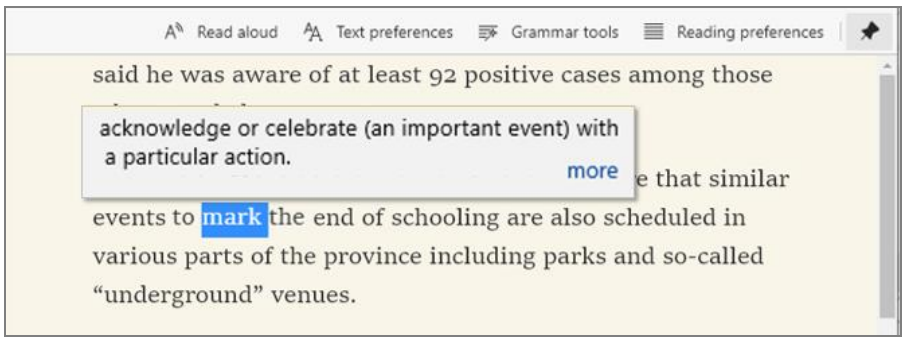

Figure 7: Ideal content of pop-up window to replace the one shown in Figure 5

The ideal solution would be a program that could detect the concrete meaning of any word occurring in a text and, at the same time, only upload the minimum of data required to assist the reader. Figure 7 shows how such a solution

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

could look like, based on the definition marked by the dotted circle in Figure 6. This type of solution is not viable yet, but technology is taken us closer and closer to it. It is now possible to assign a word occurring in a text to a specific sense with a probability of up to 90 percent. The technique makes use of machine learning, big corpora, semantic annotation, and lexicographical databases with a large number of words, senses, and example sentences. Although available, it requires extremely big processing power, and for the time being, only companies like Google can afford to acquire the necessary hardware. To our knowledge, it has not yet been applied to dictionaries integrated into e-readers. But it is under continuous development, and as has happened with other technologies, we can expect it to become cheaper and more efficient within a short span of years. It would probably not be a bad idea that all relevant stakeholders, among them lexicographers and L2 experts, start reflecting and preparing themselves for this Brave New World.

4. Dictionaries integrated into learning apps

The proposed content of the pop-up window shown in Figure 7 is inspired by Huang and Tarp (2021), who suggested a similar solution for an L2-learning app for Chinese learners of English. The two authors based their proposal on a significant difference between learning apps and other tools used to assist the reading of books and Internet texts. The number of texts and words appearing in the latter is literally speaking *infinite*, whereas learning apps include a relatively *limited amount* of texts and words. This difference allows other methods to be applied. The required technology is already available and relatively simple. Instead of artificial intelligence which was a precondition for the proposal in the previous paragraph, the two authors recommend *human-assisted intelligence* based on an interdisciplinary collaboration between language experts, lexicographers, and information engineers.

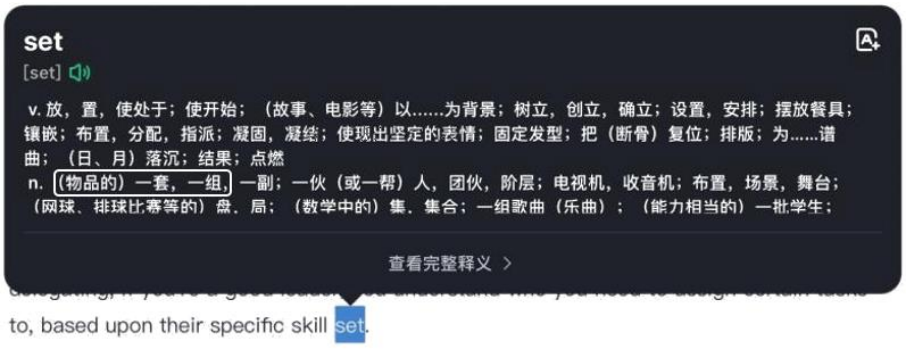

Figure 8: Pop-up window with *set* in Kaiyan app

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

To illustrate their proposal, Huang and Tarp (2021) discuss an example from the *Kaiyan OpenLanguage* learning app where a user clicks on the word *set* in the text (see Figure 8). The displayed pop-up window contains a big majority of lexicographical data that are completely irrelevant in the concrete context. More than half of the data belong to the verb *set*, although the word clicked on was a noun. In fact, only the characters inside the white frame relate to the concrete meaning of *set*, i.e. less than 10 percent of the total. This meaning item is not easy to find at first glance. It will probably take several seconds to detect, evaluate, and choose the right meaning of *set*. The large amount of irrelevant data obstructs the information search process. Huang and Tarp (2021) therefore suggest an alternative solution (see Figure 9) and explain how it can be obtained with current technology.

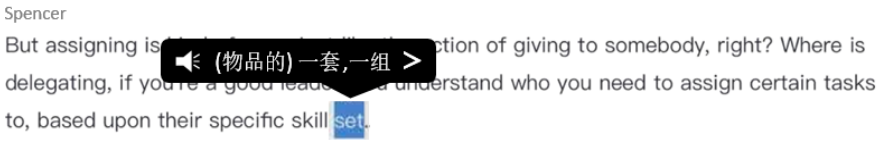

Figure 9: Ideal content of pop-up window to replace the one shown in Figure 8

The difference between the proposed pop-up window and the one currently used in the *Kaiyan OpenLanguage* app becomes crystal-clear if we compare Figures 8 and 9. The proposal contains an absolute minimum of items to meet the users' needs in the concrete context. Huang and Tarp (2021: 87) explain the underpinning philosophy:

> The main idea is that the pop-up window should only include items that can be justified by the immediate user needs. Thus, it merely consists of a speaker icon, a meaning discriminator followed by two equivalents, and a signifier (>). The central item is the definition (or equivalents) that directly assists understanding of the course text. […] The window also includes a speaker icon to service learners who, as recommended by language didactics, read aloud and may need to listen to some of the words to pronounce them right. Finally, it provides a widely used signifier that affords access to the whole article and is well-known to most netizens.

As can be seen, the proposed default pop-up window breaks with well-established features of the traditional dictionary article, as it does not contain lemma, part of speech, inflectional morphology, other senses, or any other lexicographical data. All these items are considered irrelevant as an immediate response to the specific information need, that is, to understand the text. The proposal is rooted in a millenarian cultural practice.

The traditional dictionary is the result of a long historical development. It started, both in China and Europe, in ancient times when the old scribes copied

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

manuscript works from earlier periods and inserted glosses to explain obsolete and difficult words; see McArthur (1986), Stathi (2006), and Yong and Peng (2008). These glosses were later compiled into glossaries which, over the centuries, developed into the modern dictionary. The process had both advantages and disadvantages and saw at least three important innovations. The first was the invention of the lemma that assigned all grammatical forms of nouns, verbs, and adjectives to a single and generally accepted canonical form. The second was the macrostructure that organises all words treated in the dictionary. In the beginning, it was systematic with the words arranged in the same order as they appeared in a specific book or text. Later, other structuring criteria became dominant. In Europe, the preferred criterion was the alphabet. It started with the first letter of each word, then came the second and the third, and so on, until the macrostructure finally appeared strictly alphabetic. The third invention was the microstructure that became increasingly complex when social and economic development required more and more lexicographical data to be included in the dictionary articles, which also led to a condensed, unnatural description language and the use of codes and abbreviations.

The positive outcome of this long historical process was a lexicographical product (the dictionary) that could be widely consulted in different contexts and not only in connection with a specific text or book. The negative outcome was that the introduction of the lemma and the alphabetic macrostructure required a complex mental process from the users, whereas the microstructure presupposed that the latter developed still better "reference skills".

The solutions proposed in Figures 7 and 9 represent an attempt to build upon the positive aspects and avoid the negative ones detected in the history of lexicography. The default pop-up windows have several advantages:

— The overall design follows the principles of human-centered, or user-centered, design, as recommended by Tarp and Gouws (2020).
— The windows can be activated and used intuitively and do not require special instructions or skills, as recommended by Rundell (2015).
— The response to learners' needs is immediate and represents an example of good communication, as recommended by Norman (2013).
— The lexicographical data are contextualized and provided directly in the context where an information need occurs, as recommended by Tarp and Gouws (2019).
— The windows contain only the required minimum of data and, in this way, avoid negative phenomena such as data and information overload, as recommended by Gouws and Tarp (2017).

All these design features guarantee that the consultation process does not disturb the learners' reading flow and focus on the text, thus creating the optimal conditions for incidental learning as defined above. What is more, the design also provides easy and intuitive access to additional data by clicking on the respective signifiers ("more" and ">"). In this way, the learners can switch to

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

intentional learning whenever they need it for one reason or another. This step allows them to fully benefit from all the positive aspects of the modern dictionary, provided that the presentation of the additional lexicographical data also follows the mentioned principles of user-centered design.

5. Multi-word units of meaning

The translation of the Chinese characters used to explain the meaning of *set* in Figure 9 is a "suite or series, group (of things)". This short definition of *set*, as it appears in the text, is sufficient to make it understandable to the learner. But it is not completely satisfactory. In the specific context, *set* is part of the frequent word combination *skill set*, which represents a so-called extended unit of meaning. In a posthumous article, Sinclair (2010: 37) discusses the lexicographical treatment of this type of multi-word combinations and recommends their lemmatization:

> The evidence from corpora adds up to a strong case for extending the treatment of multi-word units of meaning — a much wider concept than idiom — and giving them the same status as the usual headword.

Sinclair's reflections are also relevant to e-reading tools. Learners cannot be expected to recognise multi-word units of meaning when they meet them in a text. If they do not understand them, they will tend to click on the individual words — in the above case, either *skill* or *set*, or both of them separately. The misinterpretation may derail the consultation process with negative consequences for the reading flow. Designers of e-reading tools should therefore make provisions for this challenge. The tools should be designed to give a lexicographical response that covers the multi-word units of meaning as a whole when users click on one of their component parts. Huang and Tarp (2021) have shown how it could be done in learning apps combining good handicraft and relatively simple programming. But the suggested method presupposes a limited amount of texts and is not an option in e-readers and similar tools where the texts, in theory, are unlimited (see above). Fortunately, current technology makes allowance for another method that is already used to analyse texts. To illustrate how it works, we have taken the following sentence from the article used in Paragraph 3 (Singh 2020):

> Zikalala appealed to parents and pupils not to organise or take part in celebrations in the province which flout current Covid-19 safety protocols and endanger lives.

If readers, especially non-native speakers of English, have problems understanding this sentence, they may click on one or more words, for instance *part*, to get lexicographical assistance. The underlying program then automatically starts exploring the surrounding words to detect extended units of meaning.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

Figure 10 is a schematic representation of the process that takes place and only lasts a few nanoseconds. It goes more or less like this: The program starts looking at the first word after *part* to see if there is a recognisable multi-word unit; it then continues with the first word before *part*, the second word after *part*, the second word before *part*, and so on. In the example shown in Figure 10, it examines the six words closest to *part*, but it could be programmed to do more. In this way, it can detect extended units of meaning consisting of two or more words, even if they are separated from each other by a few other words.

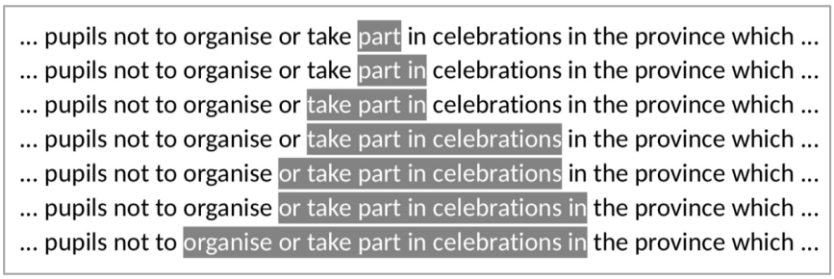

Figure 10: Schematic representation of technique to detect multi-word units

In the sentence from Times Live, *take part* is a multi-word unit with its own specific meaning (*participate*). Hence, when readers are unaware of this and click on *part*, the described technique makes it possible to give them a lexicographical response to *take part*. The instantaneous and context-adapted response is likely to boost their learning of this and similar multi-word units of meaning. The precondition, however, is that the underlying program already "knows" these units and can react upon them. This underscores the relevance of their lexicographical treatment and lemmatisation as recommended by Sinclair (2010). As mentioned above, the described technique is already available but, as far as we are informed, it has not yet been applied to dictionaries integrated into e-reading tools. Its application does not only depend on information engineers. It also requires that lexicographers compile the high-quality lexicographical databases that allow cutting-edge programming to prosper for the benefit of L2 learners, whether engaging in incidental or intentional learning.

6. Digital writing assistants

In a recent article, Graham (2020: 535) contends that:

> the sciences of reading and writing are too narrowly focused on how to teach either reading or writing and not focused enough on how these two skills can be used to support each other.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

Graham, therefore, recommends that the two sciences become "more fully integrated". From the perspective of lexicography, these statements seem both logical and relevant for the topic under discussion. Just like reading, writing is increasingly performed on digital devices. Together with oral text production, it is through writing that learners activate their L2 vocabulary and train spelling, morphology, and syntax. Similar to what happened in Figures 2 and 3, writing can also be a source of incidental learning. Where the former was human-assisted, incidental learning in connection with writing must be machine-assisted.

Digital writing assistants have some interesting possibilities in this respect, especially those integrated into the text-processing programs learners typically use when they write. When these tools, for instance, are installed on smartphones and tablets and indicate the most likely word terminations, they may stimulate spelling; see, e.g., the one presented by Tarp et al. (2017). When they suggest the most likely words to follow in the sentence, they guide their users into the exciting world of word combinations, among them the ones discussed in the previous paragraph. In this respect, Hanks (2013: 399) distinguishes between *possible* and *probable* combinations and observes that "the number of probable combinations … is rather limited", although "the number of possible combinations may in principle be limitless". Rundell (2018: 6) adds:

> Although corpus analysis enables us to observe the inbuilt predictability of most language output, much of this is far from predictable to a learner …

The desired predictability can be boosted by some of the techniques applied in writing assistants. However, it is not sufficient to look forward based upon the words already typed. It is also necessary to look back on these words to check whether one or more of them has to be changed. This is the advantage of the technology applied in writing assistants like Grammarly, LanguageTool, and ProWritingAid. These tools do no predict the next word in the sentence, but they come up with alerts and suggestions only a few seconds after typing the words. So far, the technology only seems to handle word combinations to a certain extent and, thus, still needs to be improved in this regard. It is, however, increasingly efficient and convincing in other aspects like word choice, spelling, morphosyntax, and punctuation. The suggestions refer to parameters like correctness, clarity, conciseness, and conventions, and include both error correction and text improvement marked with different colors. When users click on the marked words, a small window pops up with a brief explanation and an alternative solution, which they can insert into the text with another click. Figure 11 shows the content of the pop-up window activated after replacing *click* with *clicks* in the previous sentence.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

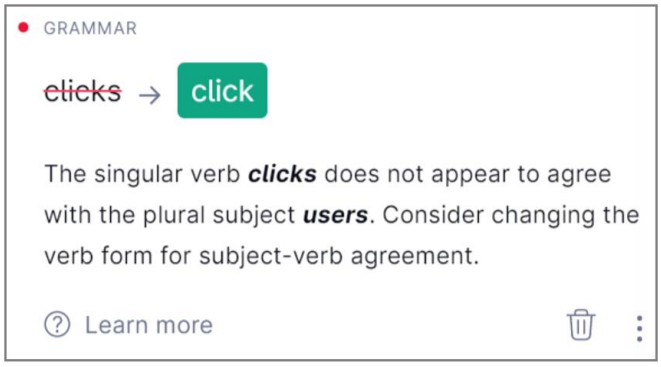

Figure 11: Pop-up window activated by clicking on an alert in Grammarly

In his visionary reflections on dictionaries, Sweet (1899: 139) also had some important recommendations concerning the treatment of grammar:

> A thoroughly useful dictionary ought, besides, to give information on various grammatical details, which, though they fall under general rules of grammar, are too numerous or too arbitrary and complicated to be treated of in detail in any but a full reference-grammar: such a dictionary ought to give full information about those grammatical constructions which characterize individual words, and cannot be deduced with certainty and ease from a simple grammatical rule. (Sweet 1899: 139)

The discussion raised by Sweet (1899) has been going on since then. Lexicographers have defended different positions on the relationship between grammar books and dictionaries and have proposed various principles for the inclusion of grammatical data and how to treat them; see Jackson (1985), Mugdan (1984), Cowie (1987, 1989), Herbst (1989), Rundell (1998), Bogaards and Kloot (2001), among many others. Agreement has not been reached, although the general tendency is to introduce still more grammatical data explained in plain language without unnecessary grammatical codes and abbreviations. It is, therefore, interesting to observe how Grammarly seems to share, at least partially, the above vision expressed by Sweet (1899). If the writers click on the signifier "Learn more", they will get immediate access to an extended usage note that explains the relevant grammatical problem in greater detail (see Figure 12). In this way, the general grammatical rule materialises in a mini-rule assigned to an individual word. It can be argued that the solution is not completely individualised and, thus, successful, but this is a case for further analysis and improvement.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

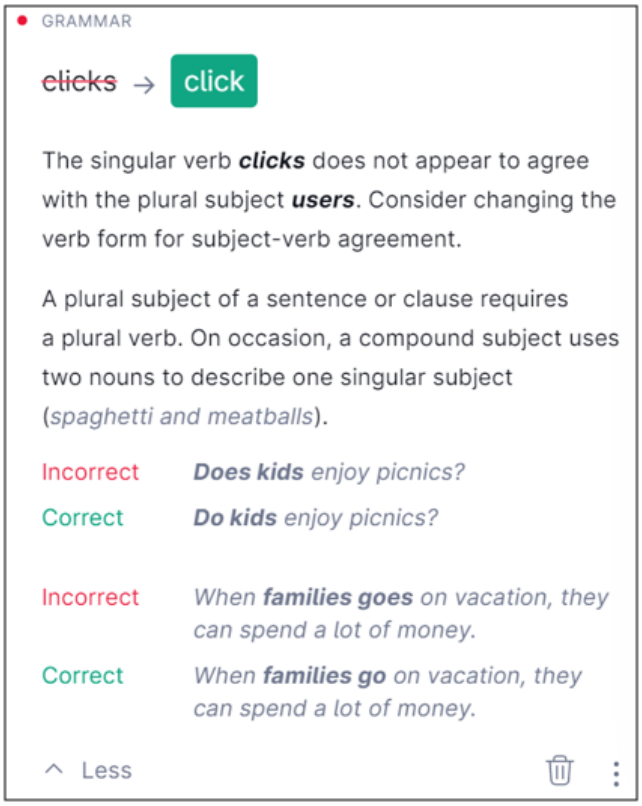

Figure 12: Grammar rule activated by clicking on "learn more" in Figure 10

The almost instantaneous suggestions and the option to insert the suggested corrections directly in the text with a simple click may pave the way for incidental learning, in so far as the production of a correct text, and not the intent to learn, is the writer's primary goal. Even so, there is undoubtedly a delicate balance between writing flow and provision of information, between focus on the text and focus on the consultation. But when the user consciously decides to click on "Learn more", it is definitely a case of intentional learning.

In Figure 3, we saw a bilingual approach to incidental learning in connection with oral communication when an L2 learner switched to his native language to ask for an L2 word he did not know or remember. Now we will demonstrate something similar in connection with written communication. We will use an example from Write Assistant (see Figure 13) discussed by Fuertes-Olivera and Tarp (2020). A Spanish learner is writing a text in English, and when he lacks an English word to express a specific idea, he types the Spanish word *cerrado* instead. The software is designed to be context-aware and automatically displays several equivalents with the ones that are most likely in the

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

concrete context listed first. This prioritised list has two functions. It can function as a reminder if the learner already knows the word but has just forgotten it. In this case, he can simply click on the appropriate equivalent to introduce it directly into the text. If he does not know one of the equivalents (in this case *sealed*), he can mark it and click on the arrow to activate a pop-up window with short L1 definitions of its various senses written in his mother tongue. If one of these meets his expectations, he can click on *sealed* to insert it into the text. The problem is solved, and he can continue writing. But if he, for one reason or another, wants to know more, he can click on one of the arrows in the pop-up window to access more lexicographical data related to the specific sense, as explained in detail by Fuertes-Olivera and Tarp (2020). If he chooses to do this, it is once more a case of intentional learning. By contrast, if the short definitions in the pop-up window are sufficient to meet his concrete needs, it may allow the learning of one of the senses of *seal*. This learning is, by definition, incidental, as long as the writer's direct and immediate intention is not to learn but to express something through a written text.

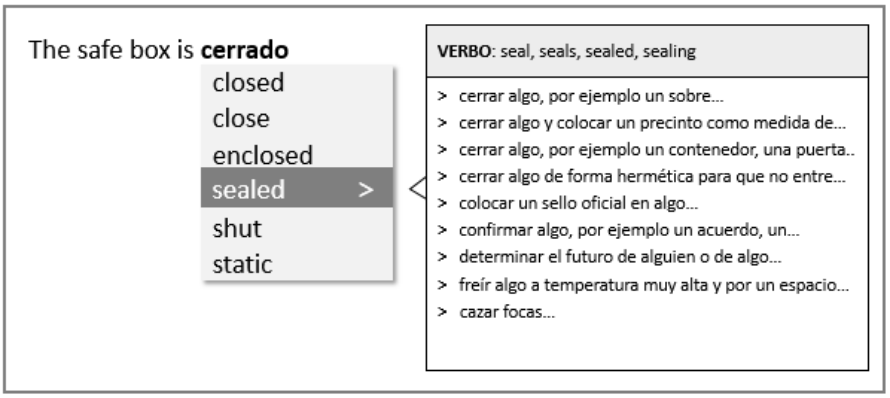

Figure 13: Prioritised equivalents and pop-up window in Write Assistant

7. Conclusions

From a lexicographical perspective, incidental learning presupposes instantaneous, contextualized, and unobtrusive assistance with an absolute minimum of lexicographical data, whereas intentional learning requires easy access to relevant additional data. The discussion above shows that it is far more complicated to create the conditions for the former than the latter. In both cases, the lexicographical data must be high quality and presented in a user-centered design that allows intuitive use. But incidental learning, as defined in Paragraph 2, also implies that the software has been trained to exclusively present the data needed in each concrete case and discard all other data. As we have

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

seen, the techniques and technologies required are already available to a certain extent, albeit insufficiently applied to lexicography. They include artificial intelligence, human-assisted intelligence, and cutting-edge programming in general. If further developed and fully integrated into lexicography, they herald a major technological breakthrough with the advent of context-aware lexicographical products, that is, a prototype of intelligent dictionaries.

The creation and quality of future context-aware lexicographical products do not only depend on the successful application of cutting-edge technologies. It also requires that lexicographers reconsider part of their discipline. It implies, among other things, refinement of the lexicographical databases that store the pertinent data and improved design of the user interfaces that present the relevant data to the target users. As Zhang (2019) rightly asserts, the current media convergence age invites lexicographers to take an innovative approach to the compilation and publication of dictionaries and, could it be added, lexicographical products in general. The challenge is *both* to integrate various media into digital dictionaries, *and* to integrate lexicography into different types of digital devices.

It may seem that some of the issues discussed in this paper go beyond traditional lexicography or, at least, belong to the borderland between lexicography and other fields of endeavour. That may be so. In any case, a better interpretation would be that they represent virgin land which the millennial discipline needs to cultivate if it wants to meet contemporary challenges. It requires a successful symbiosis of tradition and disruption; see Tarp (2019). The suggestions to improve current e-reading tools, writing assistants, and learning apps intend to remove some of the obstacles to incidental L2 learning in connection with the reception and production of written texts. It is up to future research to determine whether or not, and to what degree, these suggestions work in practice. Until then, they remain a digital information-assisted possibility of incidental learning.

References

A. Writing assistants

Grammarly. https://www.grammarly.com/.
LanguageTool. https://languagetool.org.
ProWritingAid. https://prowritingaid.com/.
Write Assistant. https://writeassistant.com/.

B. Other literature

Bereiter, C. and M. Scardamalia. 1989. Intentional Learning as a Goal of Instruction. Resnick, L.B. (Ed.). 1989. *Knowing, Learning, and Instruction: Essays in Honor of Robert Glaser*: 361-392. Hillsdale: Lawrence Erlbaum Associates.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

Bogaards, P. and W. van der Kloot. 2001. The Use of Grammatical Information in Learners' Dictionaries. *International Journal of Lexicography* 14(2): 97-121.

Bothma, T.J.D. and R.H. Gouws. 2020. e-Dictionaries in a Network of Information Tools in the e-Environment. *Lexikos* 30: 29-56.

Bothma, T.J.D. and D.J. Prinsloo. 2013. Automated Dictionary Consultation for Text Reception: A Critical Evaluation of Lexicographic Guidance in Linked Kindle e-Dictionaries**.** *Lexikographica* 29(1): 165-198.

Brown, R., R. Waring and S. Donkaewbua. 2008. Incidental Vocabulary Acquisition from Reading, Reading-while-listening, and Listening to Stories. *Reading in a Foreign Language* 20(2): 136-163.

Cowie, A.P. 1987. Syntax, the Dictionary and the Learner's Communicative Needs. Cowie, A.P. (Ed.). 1987: *The Dictionary and the Language Learner. Papers from the Euralex Seminar at the University of Leeds, 1–3 April 1985*: 183-92. Tübingen: Niemeyer.

Cowie, A.P. 1989. Information on Syntactic Construction in the General Monolingual Dictionary. Hausmann, F.J., O. Reichmann, H.E. Wiegand and L. Zgusta (Eds.). 1989. *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography. Volume 1*: 588-592. Berlin/ New York: Walter de Gruyter.

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

Fuertes-Olivera, P.A. and S. Tarp. 2020. A Window to the Future: Proposal for a Lexicography-assisted Writing Assistant. *Lexicographica* 36: 257-286.

Gouws, R.H. and S. Tarp. 2017. Information Overload and Data Overload in Lexicography. *International Journal of Lexicography* 30(4): 389-415.

Graham, S. 2020. The Sciences of Reading and Writing Must Become More Fully Integrated. *Reading Research Quarterly* 55(S1): 535-544.

Hanks, P. 2013. *Lexical Analysis: Norms and Exploitations*. Cambridge, Mass./London: MIT Press.

Herbst, T. 1989. Grammar in Dictionaries. Tickoo, M.L. (Ed.). 1989. *Learner's Dictionaries: The State of the Art*: 94-111. Singapore: SEAMO Regional Language Centre.

Huang, F. and S. Tarp. 2021. Dictionaries Integrated into English Learning Apps: Critical Comments and Suggestions for Improvements. *Lexikos* 31: 68-92.

Hulstijn, J.H. 1989. Implicit and Incidental Second Language Learning: Experiments in the Processing of Natural and Partly Artificial Input. Dechert, H.W. and M. Raupach (Eds.). 1989. *Interlingual Processes*: 49-73. Tübingen: Gunter Narr.

Hulstijn, J.H. 2013. Incidental Learning in Second Language Acquisition. Chapelle, C.A. (Ed.). 2013. *The Encyclopedia of Applied Linguistics***:** 2632-2637. New York: Wiley-Blackwell.

Jackson, H. 1985. Grammar in the Dictionary. Ilson, R. (Ed.). 1985. *Dictionaries, Lexicography and Language Learning*: 53-59. Oxford: Pergamon Press.

Krashen, S. 1989. We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *Modern Language Journal* 73*:* 440-464.

Laufer, B. and J. Hulstijn. 2001. Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-induced Involvement. *Applied Linguistics* 22: 1-26.

Leow, R.P. and C.C. Zamora. 2017. Intentional and Incidental L2 Learning. Loewen, S. and M. Sato (Eds.). 2017. *The Routledge Handbook of Instructed Second Language Acquisition*: 33-49. New York: Routledge.

Lessing, G.E. 1747. *Der Junge Gelehrte*. Stuttgart: Reclam 1965.

http://lexikos.journals.ac.za; https://doi.org/10.5788/32-2-1703 (Article)

McArthur, T. 1986. *Worlds of Reference. Lexicography, Learning and Language from the Clay Tablet to the Computer*. Cambridge: Cambridge University Press.

Mugdan, J. 1984. Grammatik im Wörterbuch: Wortbildung. Wiegand, H.E. (Ed.). 1984. *Studien zur neuhochdeutschen Lexikographie IV*: 237-308. Hildesheim/Zürich/New York: Olms.

Nagy, W.E., P.A. Herman and R.C. Anderson. 1985. Learning Words from Context. *Reading Research Quarterly* 20(2): 233-253.

Norman, D. 2013. *The Design of Everyday Things*. New York: Basic Books.

Rundell, M. 1998. Recent Trends in English Pedagogical Lexicography. *International Journal of Lexicography* 11(4): 315-342.

Rundell, M. 2015. Review Article: Shigeru Yamada. Oxford Guide to the Practical Usage of English Monolingual Learners' Dictionaries: Effective Ways of Teaching Dictionary Use in the English Class. *Kernerman Dictionary News* 23: 26-27.

Rundell, M. 2018. Searching for Extended Units of Meaning — and What To Do When You Find Them. *Lexicography — Journal of Asialex* 5(1): 5-21.

Shu, H., R.C. Anderson and H. Zhang. 1995. Incidental Learning of Word Meanings While Reading: A Chinese and American Cross-cultural Study. *Reading Research Quarterly* 30(1): 76-95.

Sinclair, J.M. 2010. Defining the Definiendum. De Schryver, G.-M. (Ed.). 2010. *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks*: 37-47. Kampala: Menha Publishers.

Singh, O. 2020. KZN Premier Lays Down the Law Against Those Attending 'Superspreader' Events. *Times Live*, 8 December 2020.

Stathi, E. 2006. Greek Lexicography, Classical. Brown, K. (Ed.). 2006. *Encyclopedia of Language and Linguistics. Vol. 5*: 145-146. Second Edition. Oxford: Elsevier.

Sweet, H. 1899. *The Practical Study of Languages: A Guide for Teachers and Learners*. London: J.M. Dent. Reprinted in 1964 by Oxford University Press.

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. 2019. Connecting the Dots: Tradition and Disruption in Lexicography. *Lexikos* 29: 224-249.

Tarp, S., K. Fisker and P. Sepstrup. 2017. L2 Writing Assistants and Context-Aware Dictionaries: New Challenges to Lexicography. *Lexikos* 27: 494-521.

Tarp, S. and R.H. Gouws. 2019. Lexicographical Contextualization and Personalization: A New Perspective. *Lexikos* 29: 250-268.

Tarp, S. and R.H. Gouws. 2020: Reference Skills or Human-Centered Design: Towards a New Lexicographical Culture. *Lexikos* 30: 470-498.

Wiegand, H.E. 2013. Lexikographie und Angewandte Linguistik. *Zeitschrift für angewandte Linguistik* 58(1): 13-39.

Yong, H. and J. Peng. 2008. *Chinese Lexicography. A History from 1046 BC to AD 1911*. Oxford: Oxford University Press.

Zhang, Y. 2019. On the Innovation of Dictionary Compilation and Publication in the Context of Media Convergence. *Research on Language Strategies* 4(6): 79-89.