

e-Dictionaries in a Network of Information Tools in the e-Environment

Theo J.D. Bothma, *Department of Information Science, University of Pretoria, South Africa (theo.bothma@up.ac.za)*

and

Rufus H. Gouws, *Department of Afrikaans and Dutch, Stellenbosch University, South Africa (rhg@sun.ac.za)*

Abstract: Traditional dictionaries offer curated data to users. Users should therefore be able to find the correct data to solve their information need. However, users don't necessarily know the exact scope of lexicographic information. Dictionary articles can still demand considerable interpretation by the user to select the appropriate meaning or equivalent.

In the e-environment, users can easily navigate between different e-sources. This is especially evident on various e-book platforms, where one can link multiple dictionaries and other sources to a text or search of the internet. Internet content is obviously not curated, and providing access to such data is therefore anathema to the traditional lexicographer. A traditional dictionary is the result of an application of data pushing procedures. The online environment enables the use of data pulling procedures that give users access to both curated and non-curated data.

These issues are illustrated by means of a number of examples that show that a large number of different and disparate information sources are easily available to the user to satisfy any specific information need, and that the dictionary is one of a plethora of information sources. The information is therefore available on demand, without risking information overload.

It is argued that, when optimising a network of information tools that constitutes a comprehensive search universe, the information retrieval structure should preferably have a circular network as application domain, rather than a linear continuum.

Keywords: CONTEXTUALIZATION, DATA DISTRIBUTION STRUCTURES, INFORMATION TOOLS, SEARCH VENUES, USER INFORMATION NEEDS, NETWORK OF INFORMATION TOOLS, DATA PULLING, DATA PUSHING, INFORMATION SOURCES, SMART E-DICTIONARIES

Opsomming: e-Woordeboeke in 'n netwerk van inligtingswerktuie in die e-omgewing. Tradisionele woordeboeke bied gekeurde data aan gebruikers. Daarom behoort gebruikers die korrekte data te kan vind om 'n inligtingsbehoefte te bevredig. Gebruikers weet egter nie altyd wat die presiese bestek van die leksikografiese inligting is nie. Woordeboekartikels stel steeds eise aan die gebruiker om die gepaste betekenis of ekwivalent te kies.

In die e-omgewing kan gebruikers maklik tussen verskillende e-bronne rondbeweeg. Dit geld veral op verskillende e-boekplatforms waar baie woordeboeke en ander bronne aan 'n teks of internetsoektoeg gekoppel kan word. Internetinhoud is vanselfsprekend nie gekeur nie en om toe-

gang tot sulke data te gee, is bykans 'n gruwel vir tradisionele leksikograwe. 'n Tradisionele woordeboek is die produk van 'n toepassing van datatrekprosedures. Die aanlyn omgewing maak die gebruik van datastootprosedures moontlik wat aan gebruikers toegang tot sowel gekeurde as ongekeurde data gee.

Hierdie kwessies word geïllustreer aan die hand van 'n aantal voorbeelde wat wys hoe talle verskillende en uiteenlopende inligtingsbronne maklik beskikbaar is vir gebruikers om enige spesifieke inligtingsbehoefte te bevredig en dat 'n woordeboek een van baie inligtingsbronne is. Die inligting is op aanvraag beskikbaar sonder die gevaar van inligtingsoorlading.

Daar word aangevoer dat wanneer 'n netwerk van inligtingswerktuie optimaal saamgestel word wat 'n omvattende soekuniversum bied, moet die inligtingsonttrekkingstruktuur verkieslik 'n sirkel-netwerk eerder as 'n lineêre kontinuum as toepassingsgebied hê.

Sleutelwoorde: DATA STOOT, DATA TREK, DATAVERSPREIDINGSTRUKTURE, GEBRUIKERSINLIGTINGSBEHOEFTE, INLIGTINGSBRONNE, INLIGTINGSWERKTUIE, KONTEKSTUALISERING, NETWERK VAN INLIGTINGSWERKTUIE, SLIM e-WOORDEBOEKE, SOEKPLEKKE

1. Introduction

Some dictionaries, especially printed dictionaries, can still be regarded as stand-alone reference products. The emergence and establishment of online lexicography did not only change the medium of dictionaries, their structures, contents and presentation of data. It has had a far-reaching influence on the nature and extent of lexicography as a field of research and on the position of dictionaries as reference sources. Successful utilisation of possibilities available in the online environment demanded an increase in interdisciplinary collaboration. In this regard the importance of mutual research efforts in the fields of lexicography and information science has been apparent for some time, as is evident from a number of articles in Fuertes-Olivera and Bergenholtz (2011) and Granger and Paquot (2012). More recently, the results of these joint efforts can be seen in various publications, e.g. Ball and Bothma (2018), Bergenholtz and Agerbo (2017), Bergenholtz, Bothma and Gouws (2015), Bothma (2018), Bothma and Bergenholtz (2013), Bothma, Gouws and Prinsloo (2016), Bothma, Prinsloo and Heid (2018), Bothma and Tarp (2014), Tarp and Gouws (2019).

One of the realities of online referencing is that dictionaries function along with numerous other reference sources in a network of information tools. This has been stated succinctly by Varantola already in 2002:

Dictionaries need not be regarded as stand-alone lexical tools that should provide all the answers that the users need about language in use. In my vision, the future dictionary is rather an integrated tool or a number of tools in a professional user's toolbox where it coexists with other language technology products such as encyclopedic sources of reference, different types of corpora, corpus analysis tools (Varantola 2002: 34-35).

Successfully accessing and using such sources pose definite challenges to users.

Lexicographers and metalexigraphers need to respond to these challenges in order to ensure the continued use of dictionaries. This paper looks at some of the issues relevant to lexicography when negotiating the position of smart e-dictionaries in a network of information tools.

2. Finding the desired item

Traditional dictionaries, both printed and online variants, typically offer curated data to users. Users should therefore be able to find the correct data to solve their information need. Lexicographers add a caveat — the information need should be a lexicographic information need¹. However, all users don't necessarily know the exact scope of lexicographic information presented in a given dictionary and whether this information will assist them to execute a successful dictionary consultation procedure. A further complication arises where words are not included in the dictionary. Even if a required lemma is found, the dictionary does not necessarily provide the user with only the correct information for the context or does not necessarily contain the required information to satisfy a specific need. To illustrate this: a user reads the phrase "like burrs in a rough cloak" and is unfamiliar with the word *burrs*. When consulting the first monolingual English dictionary available to that user, in this case *Dictionary.com*, the user finds the article that has the lemma *burr* as guiding element, as seen in Figure 1. The occurrence of *burr* as a noun is treated as a polysemous lexical item. Although none of the given paraphrases of meaning reflects the meaning of the word *burrs* with which the user was confronted, the user does not know that and might opt for an inappropriate paraphrase of meaning, e.g. that given as sense number two. This will lead to incorrect text reception.

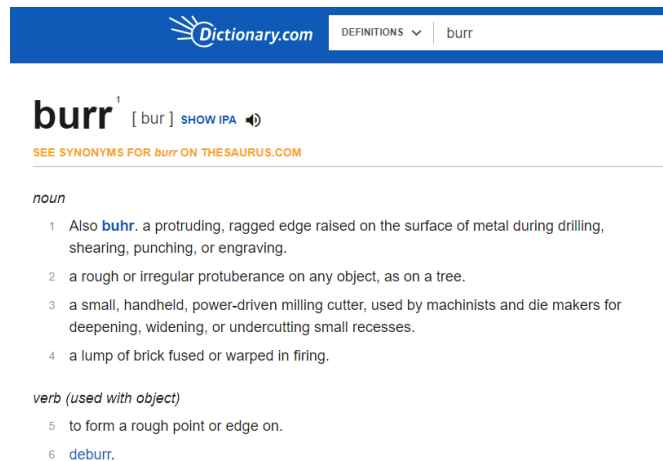


Figure 1: Partial article of the lemma sign *burr* from *Dictionary.com*

When a dictionary user has a text reception problem with the word *burrs* and consults a dictionary article with the required word as lemma it can still demand considerable interpretation by that user to realise that the items included in the article of the specific dictionary article do not provide an answer to the question that prompted the specific dictionary consultation. If the user is able to correctly interpret the information presented by the items in the dictionary article, they will realise that they have to consult another dictionary in order to find a solution to the text reception problem they have encountered. Such a second dictionary may also fail to provide the necessary assistance and even when a dictionary that does offer the required information is found, the user often still faces a real challenge to select the appropriate paraphrase of meaning. The complexity of such a search procedure is aggravated when the selected dictionary lacks sufficient cotextual items in the subcomments on semantics to ensure an optimal and unambiguous retrieval of information from the items given as paraphrases of meaning. This can be seen in Figure 2 where the user needs to read through the full dictionary article (or at least up to item 5) to come to the conclusion that the item giving the paraphrase of meaning for sense 5 of the lemma *burr* is the appropriate one:

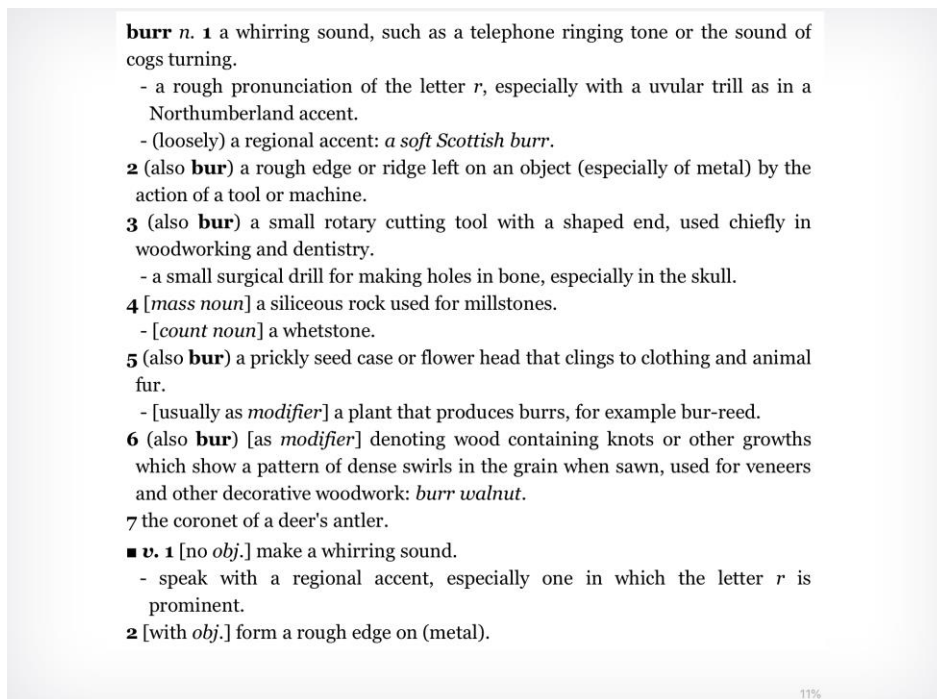


Figure 2: Partial article of the lemma sign *burr* in the *Oxford Dictionary of English* (linked Kindle edition)

Especially in a paper-based environment, users often have to consult multiple sources, which could be time intensive, and could often still lead to users' information needs not adequately satisfied.

The situation is different in the e-environment, where the user can navigate between sources — both e-dictionaries and other e-sources. This is especially evident on various e-book platforms. One can link multiple dictionaries to a text. By clicking on a word, the user has access to them but when that word is not a lemma in a specific dictionary the user can obtain misleading guidance, as in a case where a user needed an explanation for *Tamburlaine* as found in the phrase *I loved both Tamburlaines*. Clicking on that word in the text the user is directed to a lemma in the linked Kindle dictionary that closely resembles the form of the required word, i.e. the lemma *Tamerlane*, as seen in Figure 3, but such a dictionary could be completely misleading because the user has been directed to an incorrect interpretation of the word:

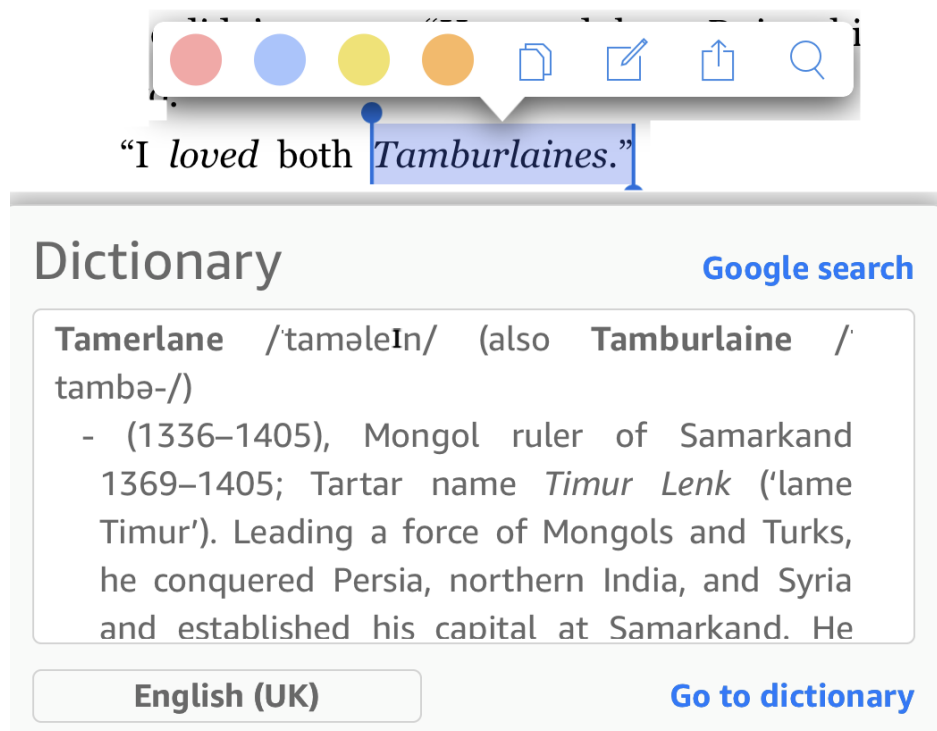



Figure 3: A word, *Tamburlaines*, in the Kindle app on an iPad, linked to the dictionary

Yet again, it is the responsibility of the user looking for help to realise that the consultation has not been successful, and the search has to continue. The user could also go directly to Google, to search the internet, resulting e.g. in a hit like the one presented in Figure 4:



Tamburlaine 

Play by Christopher Marlowe

Book preview
Full book available

[READ NOW](#)

3,6/5 · [Goodreads](#)

82% liked this book  

Google users

Tamburlaine the Great is a play in two parts by Christopher Marlowe. It is loosely based on the life of the Central Asian emperor, Timur. [Wikipedia](#)

First performance: 1587

Playwright: [Christopher Marlowe](#)

Characters: [Zenocrate](#), [Bajazeth](#), [Mycetes](#), [Cosroe](#), [Techelles](#), [MORE](#)

Figure 4: Results of a search for *Tamburlaines* on Google

In the consultations of which figures 3 and 4 reflect the results, the specific dictionary is one of a range of information tools available to the user — but not necessarily the source offering the required assistance. It is expected from users to apply their minds in order to ensure a successful information retrieval. This could imply that the user has to move between different information tools constituting a network of reference sources, that is a representation with different reference sources occurring in different positions on this representation.

3. Moving between different information tools

When reading a text on a Kindle, a reader can be compelled to employ a variety of procedures in order to obtain the meaning of a word found in a text on the Kindle. If a user e.g. clicks on the compound *coney-catcher* in a text situated in 16th century London in the UK, the user gets access to the Kindle-linked dictionary and is taken to the lemma sign *coney* — as seen in Figure 5:

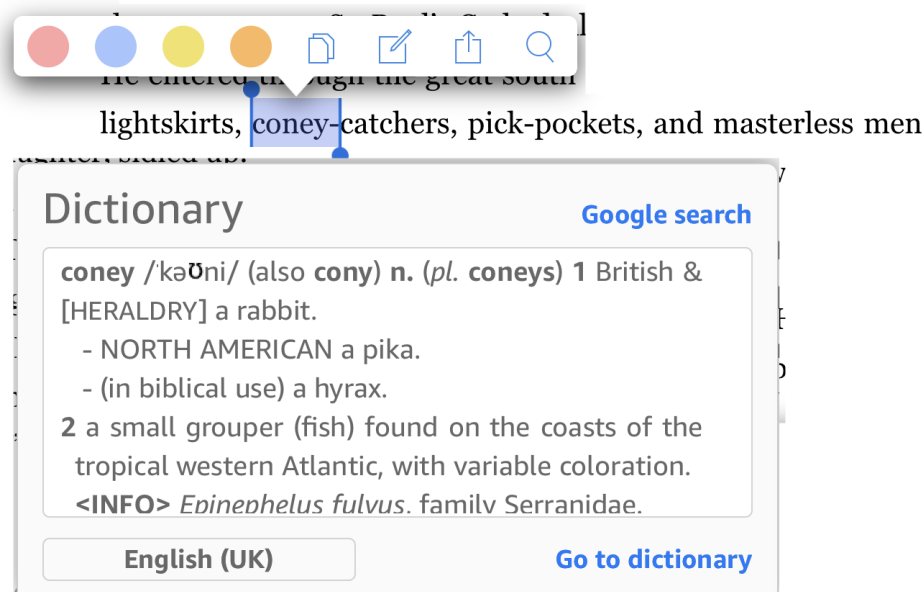


Figure 5: The word *coney* in the Kindle app on an iPad, linked to the dictionary

This dictionary does not include the compound *coney-catcher* and the user has to negotiate the treatment allocated to the simplex *coney* in order to try and solve the text reception problem. According to the Kindle-linked dictionary the lemma *coney* has two polysemous senses, referring to a "rabbit" or a "small grouper (fish) found on the coasts of the tropical western Atlantic." Within the

text in which the compound was found, situated in London, the user can rightfully deduce that the second sense does not apply. The question remains whether the first sense is the appropriate one. If it is the appropriate sense the reader therefore

- has to surmise that the compound would most probably not be the literal sense
- has to find the figurative sense, which fits perfectly to the context.

The cotextual environment of *coney-catcher* in the source text combines this compound with the lexical items *pick-pockets* and *masterless men*. How does a literal *rabbit catcher* fit into this environment? The user has to consult other sources and a Google search leads the user to various possibilities, e.g. illustrations like that given in Figure 6:



Figure 6: Coney-catcher

There are also occurrences of this word in numerous texts. From these occurrences the user might perhaps be able to deduce that a coney-catcher is a kind of swindler but it is not given. Fortunately for the traditional dictionary user, the Google search also offers reference works like *Lexico*, where the article of the lemma *coney-catcher* can be accessed, and *Wikipedia*, where the article of the lemma *coney-catching* can be accessed, as seen in Figure 7 and Figure 8 respectively:

The screenshot shows the Lexico website interface. At the top, there is a navigation bar with 'DICTIONARY', 'THESAURUS', 'GRAMMAR', 'EXPLORE', and 'SPANISH DICTIONARY'. Below this is the 'LEXICO' logo, 'Powered by OXFORD', and the text 'Oxford English and Spanish Dictionary, Thesaurus, and Spanish to English'. A search bar contains the text 'coney-catcher'. Below the search bar is a banner for 'Adobe Creative Cloud' with the text 'Make it with Creative Cloud. Apps for photography, design, video and web. Join now'. The main content area shows the breadcrumb 'Home > UK English > coney-catcher'. The title 'coney-catcher' is displayed in a large font, with a pronunciation guide '/ˈkəʊnt, kətʃə/' and social media icons for Facebook, Twitter, and LinkedIn. Below the title, the word is classified as a 'NOUN'. Two definitions are provided: 1. 'archaic, historical A swindler, a cheat, a trickster; a perpetrator of confidence tricks; a deceiver. Also figurative. Popularized by Robert Greene and remaining common for 60 or more years.' 2. 'dialect, historical A person (occasionally also a trap or a dog) that catches rabbits. Compare "rabbit catcher".'

Figure 7: Partial article of the lemma sign *coney-catcher* in *Lexico*

The screenshot shows the Wikipedia article for 'Coney-catching'. The page includes the Wikipedia logo, a navigation bar with 'Article' and 'Talk' tabs, and a search bar. The title 'Coney-catching' is prominently displayed. Below the title, it states 'From Wikipedia, the free encyclopedia'. The main text begins with 'Coney-catching is Elizabethan slang for theft through trickery. It comes from the word "coney" (sometimes spelled conny), meaning a rabbit raised for the table and thus tame.^[1]' A note indicates 'A coney-catcher was a thief or con man.^[2]' The text continues: 'It was a practice in medieval and Renaissance England in which devious people on the street would try to con or cheat vulnerable or gullible pedestrians. The term appears in *The Taming of the Shrew* and *The Merry Wives of Windsor* by William Shakespeare, and in the John Florio translation of Montaigne's essay, "Of the Cannibals." The term was first used in print by Robert Greene in a series of 1592 pamphlets,^[3] ^[4] the titles of which included "The Defence of Conny-catching," in which he argued there were worse crimes to be found among "reputable" people, and "A Disputation betweene a Hee Conny-catcher and a Shee Conny-catcher." Kirby Farrell wrote a book called *Coney-catching* in 1971. Virginia Woolf mentions "cony catchers" in her book "Orlando".' A 'References' section follows, listing four sources. A note at the bottom states '[stAb] This vocabulary-related article is a stub. You can help Wikipedia by expanding it.'

Figure 8: Partial article of the lemma sign *coney-catching* in *Wikipedia*

The treatment in these articles shows that a *coney-catcher* is a thief, a con man, a swindler or a trickster. This can also be confirmed by the treatment found in the *Oxford English Dictionary* (<https://www.oed.com/view/Entry/40917>).

Finding content in a source that is not a traditional dictionary leads to another challenge for the user. Such content is not curated, and providing access to such data is therefore anathema to the traditional lexicographer, because they have no control over the data presented to the user. Venturing into domains where both curated and non-curated content can be found, demands a different approach to the information quest by the user, i.e. to complement the search results provided by a traditional data pushing model with that of a data pulling model.

4. Data pushing and pulling

Research in various fields, including information science and lexicography, has been directed at aspects of data pulling and data pushing, cf. among other Duan, Gopalan and Dong (2005), Deolasee et al. (2010), Müller-Spitzer (2013) and Gouws (2018). Duan, Gopalan and Dong (2005: 2-3) formulate the difference between a push and a pull medium as follows:

In the sender-push model, the sender knows the identity of a receiver in advance and pushes the message in an asynchronous manner to the receiver. The receiver accepts the entire message, may choose to optionally examine the message, and then accept or discard it. An important aspect of the sender-push model is that the entire message is received before any receiver-side processing is performed.

In the receiver-pull model it is the receiver who initiates the message transfer by explicitly contacting the sender. The sender passively waits for the receiver and delivers the entire content upon receiving a request. Since it is the receiver who initiates the message transfer, the receiver would have explicit greater control over the message transfer and implicit greater trust in the received content, than in the sender-push model.

With regard to this distinction in the field of lexicography Müller-Spitzer (2013: 369) argues:

Generally, the Internet is considered to be a 'pull-medium' rather than a 'push-medium' like television, radio, or books ... Therefore, users are both sender and receiver. They are active in 'pulling' data from the website, saving relevant parts, etc. Thus, the Internet provides a very new form of communication in general.

And further:

It is communication in an innovative combination with new media.... This general property of the Internet as communication medium obviously has consequences for the property of online dictionaries as one type of text on the Internet. The process of pulling and, thus, representing lexicographic data according to a

user request is essential for EDs and must be considered when the textual structures of EDs are being looked at. (ED = electronic dictionary)

Both pushing and pulling approaches are relevant in lexicography. The need for the traditional dictionary still remains, but dictionaries do not necessarily have to be the final destination of an information search. Lexicographers need to plan and structure their e-dictionaries in such a way that they could be either a final destination or a transit area for information seekers. With dictionaries as a point of departure users should be able to obtain data from other sources. This will demand the use of both pushing and pulling approaches. For lexicographic applications of a pulling approach, Gouws (2018: 13) has introduced the notion of a lexicographic data pulling structure. He defines it as follows:

A lexicographic data pulling structure can be regarded as a structure consisting of a number of ordered elements that establishes the steps a dictionary user can follow in order to access from a given position in an existing online dictionary (where this position could be an item or search zone in a dictionary article or an article-external position, e.g. an entry in an outer feature) to dictionary-external sources from which the user can retrieve information to satisfy a specific lexicographic need.

Where traditional printed and online dictionaries result from the application of data pushing procedures, the online environment in addition enables the use of data pulling procedures that give users access to both curated and non-curated data. Deolasee et al. (2010) discuss a different kind of environment but their argument in favour of the need for combining push and pull to disseminate dynamic data also applies when utilising a network of information tools. Although a dictionary can still be used as an isolated stand-alone reference source, it should also be seen as integrated into a range of other sources where it populates a position in a network of sources. Yet again statements by Deolasee et al. (2010), albeit with regard to a different environment, also apply to this environment when they say:

A pull-based approach does not offer high fidelity ... A push-based algorithm can offer high fidelity ...

The occurrence of both dictionaries and other reference sources in a network of information tools demands that users have to be educated to distinguish between different data acquiring approaches and to be fully aware of the risks and opportunities when accessing non-curated data.

5. Reaching data

Regarding a dictionary as one of the sources in a network of information tools implies that users looking for information do not necessarily have to consult the dictionary, but they could also go directly to one or more of the other

sources on this network, albeit that these sources may contain non-curated data. When it comes to the use of dictionaries the approach in this paper supports the idea of a dictionary as a source in the information network that either supplies the user with the required information or allows access to other reference sources in a network where information retrieval results from either data pushing or data pulling procedures. This arrival-and-departure-halls status of a dictionary will not be discussed any further in this paper. The focus will rather be on various aspects of the network of information tools and the kind of reference skills and assessment the users need to ensure an optimal retrieval of information.

When employing data pushing and pulling procedures to access data in any of the sources on the information network, various methods could be used. These include direct access, moving through a portal, data on demand or bi-directional procedures between different sources. Figure 9 shows methods of accessing data — either in a dictionary (curated data) or on the web (non-curated data) from an Android mobile phone portal, whereas an iPad portal enables the result seen in Figure 10:

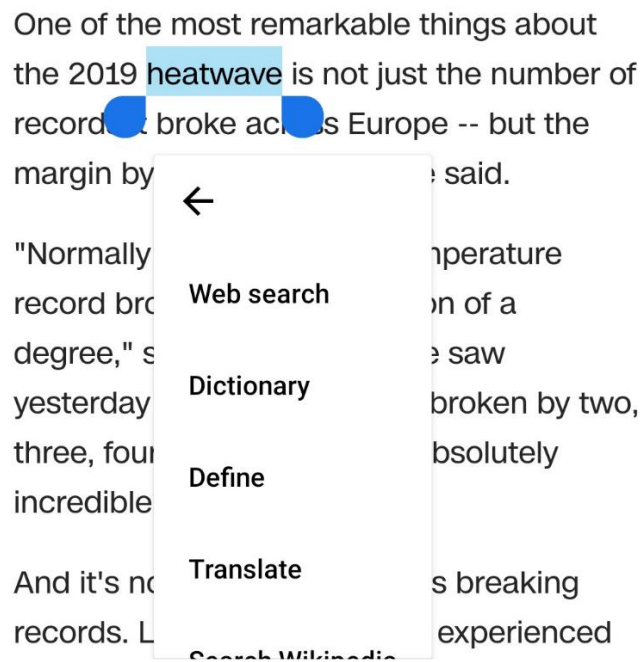


Figure 9: Pop-up portal on an Android mobile phone when the word *heatwave* is selected

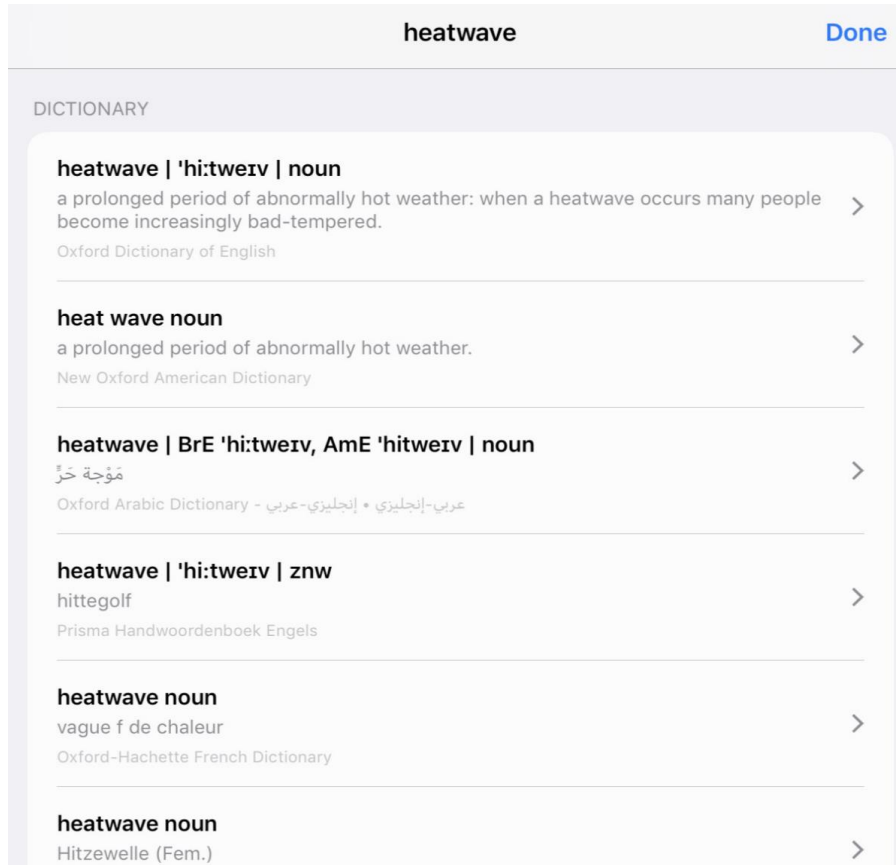


Figure 10: Pop-up portal of dictionaries on an iPad when the word *heatwave* is selected in the same text as in Figure 9

Figures 11 and 12 illustrate the information sources for searches performed in a Kindle general portal and a Kindle dictionary portal respectively.

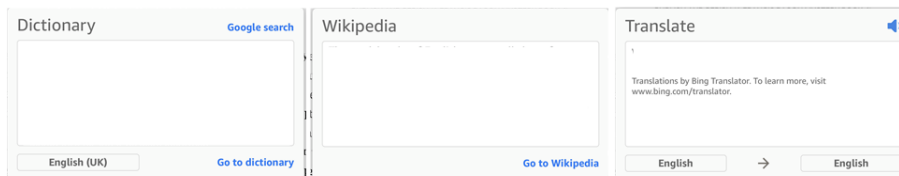


Figure 11: General pop-up portal on a Kindle, which provides access to a dictionary, Google, Wikipedia and a translation tool

In Figure 11, intentionally no content is displayed in the different results blocks.

For the translation tool, the system usually recognises the foreign language of the text, and then provides a translation. Both the source and target languages can be modified to any of the languages supported by the tool; also see Figure 17. English (UK) is the dictionary selected in Figure 11; this can be changed, as illustrated in Figure 12.

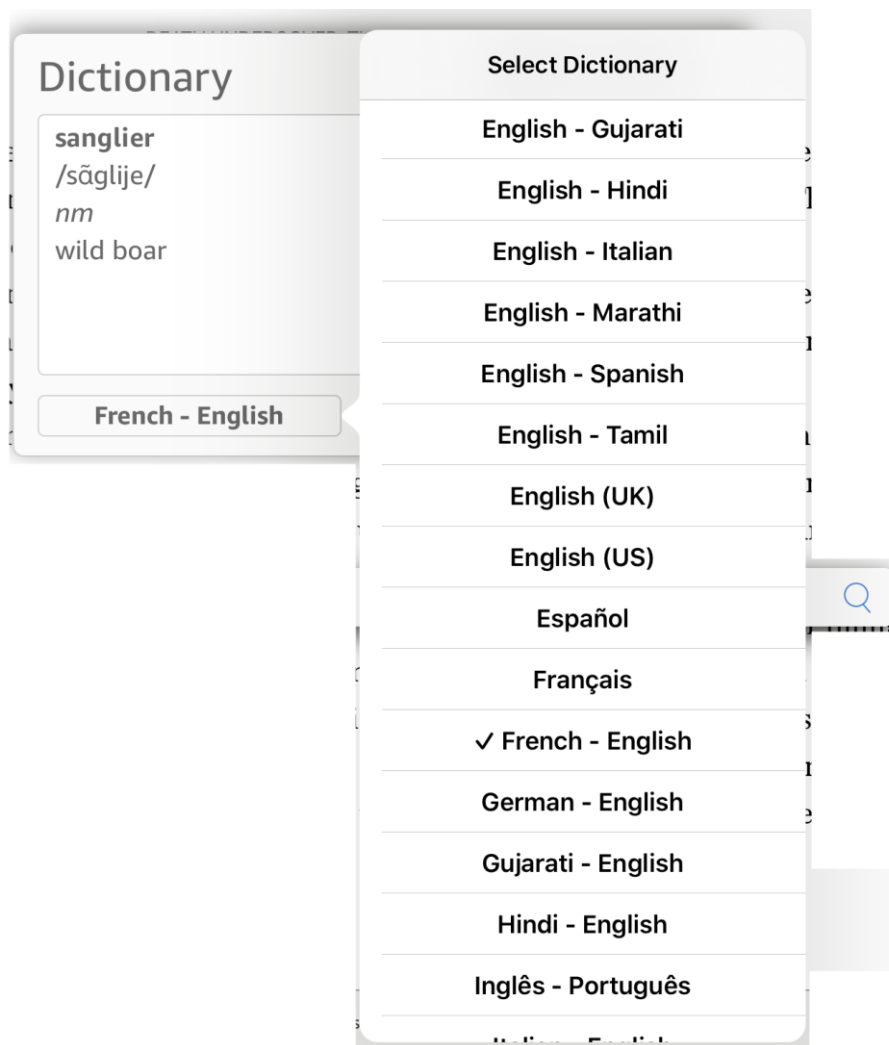


Figure 12: Kindle dictionary portal — the word *sanglier* was selected in the text (not visible on the screen capture) and *inter alia* the listed dictionaries are available

In Figure 12, a foreign language word was selected in the Kindle text. The system recognised *sanglier* as French, and used the French-English translation dictionary to provide the translation equivalent, *wild boar*. The dictionary selection can be changed to any of the free dictionaries available on the Kindle; the dictionaries need to be downloaded for use. Monolingual as well as bilingual translation dictionaries are available. A selection of these dictionaries is listed in Figure 12.

Directly accessing a dictionary article ensures curated data even though the user is not always sure whether this article delivers the solution to satisfy the lexicographic need; cf. also Tarp and Gouws (2019). The search ending in the article presented in Figure 13 might have found a satisfactory solution.

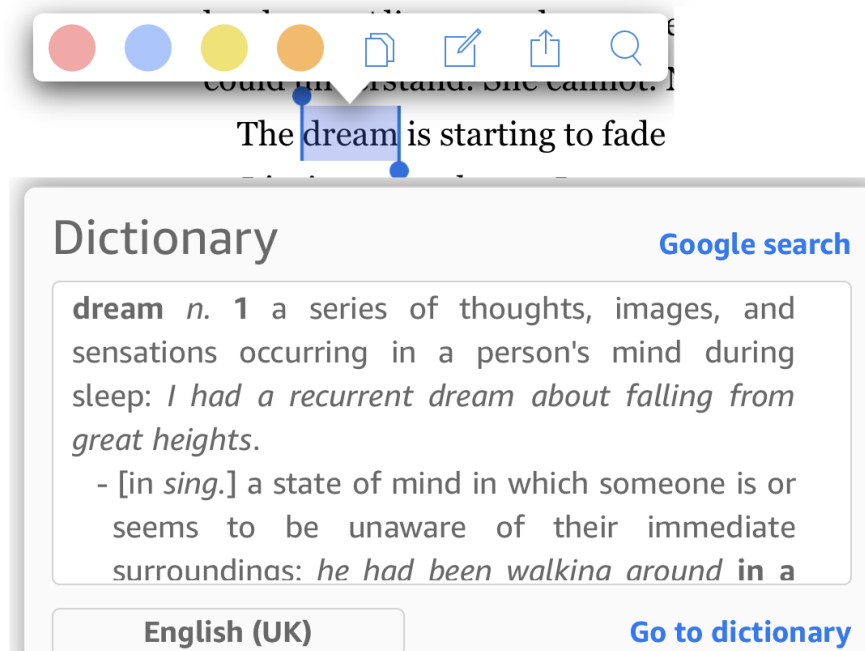


Figure 13: The word *dream* in the Kindle app on an iPad, linked to the dictionary

Figure 14 presents an article with a homonym of the required word as guiding element.



Figure 14: The word *tear* in the Kindle app on an iPad, linked to the dictionary

The homonym marker will make the knowledgeable user aware of the fact that they might have to move on to a subsequent article that also has the word *tear* as lemma. The user not equipped with these dictionary-using skills will not obtain the required assistance, although they have consulted a source that contains curated data. It is important that users should be familiar with the system of presentation in the given dictionary and be capable of assessing the information on offer.

Figure 15 shows that a search in a dictionary and *Wikipedia* for the meaning of the word *rowan* presents the user with appropriate help seeing that the word is treated as a monosemous lexical item.

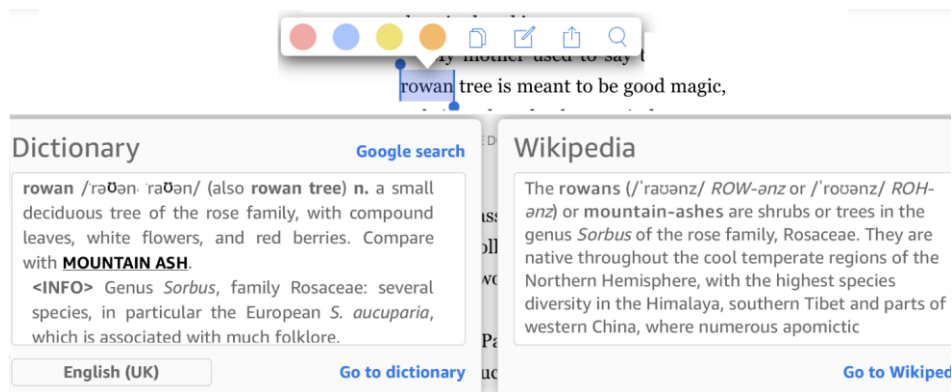


Figure 15: Clicking on *rowan* in the Kindle text provides information from the dictionary and *Wikipedia*

From the text and the two sources shown in the blocks in Figure 16, it is evident that *rowan* refers to a type of tree. The reader may want to see images of the tree, clicks on the Google link, and once the results are displayed, clicks on Images in the navigation bar. The first image is obviously incorrect, but the other images provide data that enhance the textual descriptions.

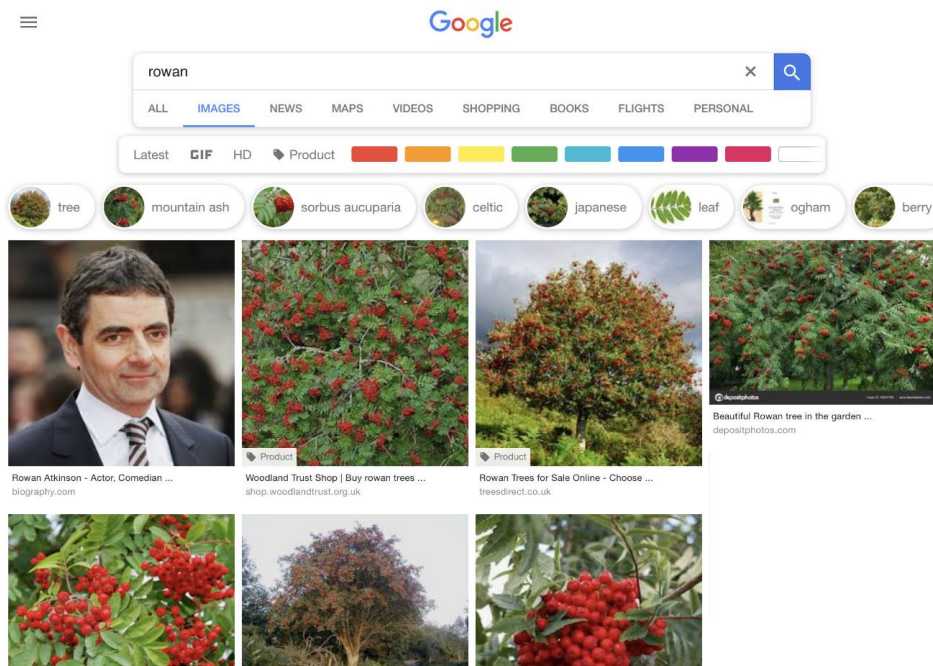


Figure 16: Image results of a search for *rowan* on Google

The value of the non-curated data on offer in some of the information tools should not be underestimated, especially if searches in these sources result from utilising data approaches — either directly or by means of a data pulling structure of a dictionary. Once the paraphrases of meaning presented in Figure 15 have convinced the user that a rowan is a tree or shrub the Google search could enhance the reference endeavour because the images give additional text reception guidance and increase the satisfaction of a cognitive need. For this success to be achieved users need to be able to apply their minds and assess the data on offer.

6. Making choices

In an online environment, users can easily move between a large number of different and disparate information sources to satisfy any specific information need. A dictionary is one of a plethora of information sources. Users have to make informed choices, based on context, prior and general world knowledge. Everywhere an information need exists, they have to make such choices. They have the option to ignore the information need by not consulting any available sources, or to explore the concept in more detail, either by consulting a dictionary article, or by delving deeper into other information sources with the dictionary as point of departure or with a dictionary-external point of departure. The information is therefore available on demand, without risking information overload.

A reader of a text on Kindle needs to obtain information regarding the form *deux chevaux*. Figure 17 shows that the linked dictionary does not provide any help but when the users opt for the translation tool, they do get some assistance. But does it solve the problem that prompted the consultation?



Figure 17: The Kindle app offers a translation of the selected words which do not occur in the linked English dictionary

To establish the meaning of *deux chevaux* in this specific context, the user needs to consult other information tools, e.g. by performing a Google search; cf. Figure 18.

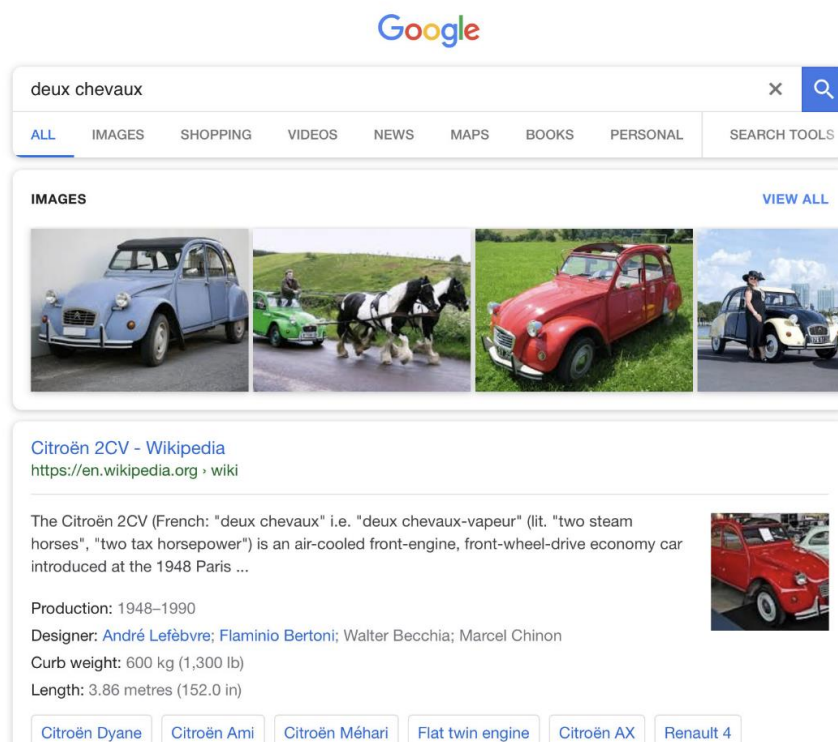


Figure 18: Images of the Citroën 2CV, commonly known as a *deux chevaux*

From the image of the Citroën in Figure 18 the required solution can be found but, yet again, the reader has to interpret the results in context.

Figures 19–21 show the additional information that can be obtained when moving between different information tools in order to find as much assistance as possible regarding a given expression.

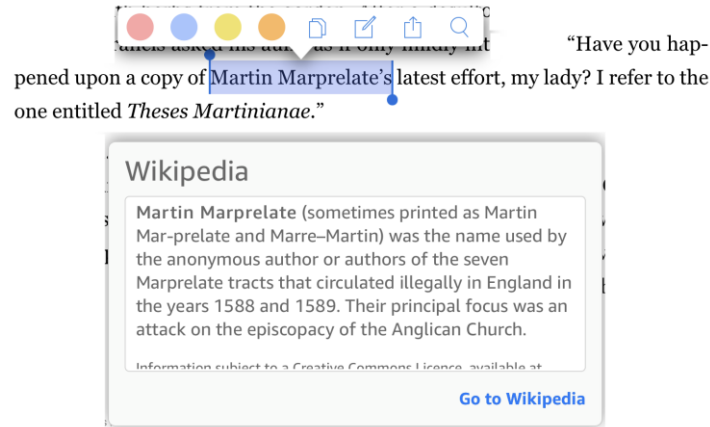


Figure 19: The linked dictionary provides no help, but *Wikipedia* provides the required information

Neither the dictionary nor *Wikipedia* provide solutions for *Theses Martinianae*. However, a Google search provides the required information (Figure 20), with an option to drill down to the annotated full text of the book (Figure 21).

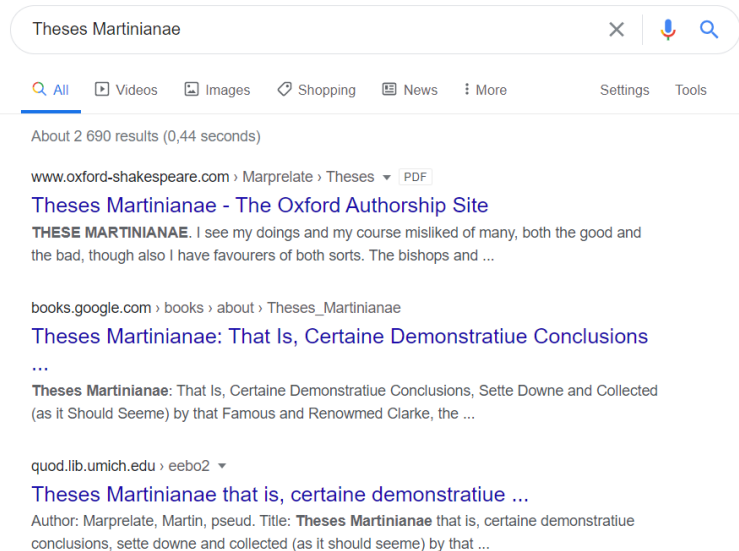


Figure 20: The linked dictionary provides no help for *Theses Martinianae*, but a search on Google provides a number of relevant links (the first few of which are shown here)

MARPRELATE TRACTS: THESES MARTINIANAE

1

THESES¹ MARTINIANAE²

That is,

Certain demonstrative³ conclusions set down and collected⁴ (as it should seem) by that famous and renowned clerk,⁵ the reverend Martin Marprelate the Great,⁶ serving as a manifest and sufficient confutation of all that ever the College of Catercaps - with their whole band of clergy priests - have, or can, bring for the defence of their ambitious and antichristian prelacy.⁷

Published and set forth as an afterbirth⁸ of the noble gentleman himself by a pretty⁹ stripling¹⁰ of his, Martin Junior,¹¹ and dedicated by him to his good neame¹² and nuncka,¹³ Master John Kankerbury.¹⁴

How the young man came by them, the reader shall understand sufficiently in the epilogue. In the meantime, whosoever can bring me acquainted with my father, I'll be bound he shall not lose his labour.¹⁵

¹ Thesis: In logic and rhetoric, a proposition laid down or stated, especially as a theme to be discussed and proved, or to be maintained against attack (in logic, sometimes as distinct from hypothesis; in rhetoric from antithesis). (OED, p.2284)

² Martin perhaps intended an allusion to Luther's 95 Theses, particularly since he calls himself 'Martin'.

³ Demonstrative: Provable by demonstration. (OED, p.517)

⁴ Although the *Theses* were printed as the fifth of the Marprelate tracts, Martin had already dealt with a number of the theses in *Hay Any Work For Cooper*, and it is possible that the entire 110 theses were 'collected' by Martin before any of the tracts were printed.

⁵ Clerk: A man of book learning, one able to read and writer, a scholar. (OED, p.348)

Figure 21: Following one of the links of Figure 20 provides an annotated full-text version of the *Theses Martinianae*

In the network of information tools, information is available on demand, and the reader can drill down to more information, if required. Information overload is therefore not a problem, as discussed in the next section.

7. Information on demand

7.1 Digital natives

Members of digital societies have become used to different possibilities to obtain information on demand. Their demand often exceeds the extent of information contained in a single source. Lexicographers need to take cognizance of this and also of the fact that a growing number of their potential target users are digital natives of which many belong to Generation Z. This generation has totally different experiences with and expectations of reference sources. They are net citizens, netizens, also known as the "Silent Generation" because of the time they spend online; cf. Parker (2013), Finch (2015). This has implications for the way in which data should be made accessible in the online environment. When looking for solutions to their problems members of the digital society seek immediate accessibility and they require uninterrupted connectivity; cf. Gouws (2017). They need to be in a position where they can pull data from a

variety of sources on the web whenever they need it and they need to be able to move from one source to the next if their need is not satisfied completely. To assist these users, guidance regarding the information on offer in a network of information tools should be accessible to them. This need has implications for the traditional notion of a data distribution structure of dictionaries.

7.2 Data distribution structures and data identification in a search universe

To ensure the best possible presentation of data in dictionaries and an optimal access to this data metalexigraphers have been concerned with an optimising of the data distribution structure of dictionaries, cf. Bergenholtz, Tarp and Wiegand (1999), Gouws (2018, 2018a). When looking at a dictionary as one source in a network of information tools lexicographers need to be aware of the data allocated to the other sources. Data distribution structures are traditionally employed in dictionaries to regulate the distribution of data in the different search venues within a single dictionary. Wiegand, Beer and Gouws (2013: 63) identify these search venues as the search field (i.e. the central list of a dictionary), the search area (i.e. each individual dictionary article) and the search zone (i.e. each slot in a dictionary article that accommodates one or more items). Gouws (2018; to appear) argues in favour of comprehensive data distribution structures, characterised by an expansion of search venues to include a search region, i.e. all the components of the textual book structure (cf. Hausmann and Wiegand 1989: 330), which will also include all components of the frame structure of a dictionary; cf. Kammerer and Wiegand (1998). For online dictionaries Gouws (to appear a) further increases the application area of the data distribution structure by identifying a search universe which covers dictionary-external sources like other dictionaries in the same dictionary portal but also other sources, including the internet. This implies that the lexicographer may put some data in dictionary-external venues, e.g. a corpus or other sources on the internet. A dictionary user venturing into a search universe from a position within a specific dictionary as point of departure, may be in a position where the retrieval of information could result from the application of both data-pushing and data-pulling procedures.

For the planning and compilation of dictionaries employed as sources in a network of information tools a comprehensive data distribution structure and extended search venues are needed. However, although a comprehensive data distribution structure can allocate data to many sources, a search universe will typically also include data of which the lexicographer of any dictionary in a network of information tools is not aware. This implies that a dictionary user accessing the search universe that constitutes the application domain of the comprehensive data distribution structure of that dictionary is likely to encounter reference sources with which they have not been familiar prior to the search. Although these data occur in the search universe these data are not positioned there by the data distribution structure of any dictionary. Where a

dictionary user may use the data distribution structure of the dictionary to identify all the available search venues and data types linked to that dictionary, a search universe allows access to additional information. Where a user executes a search for information not allocated to its position in the search universe by the data distribution structure of a dictionary occurring in a network of information tools, such a user needs guidance with regard to the available information possibilities. Knowledgeable users of such a network should be aware of what they can find where. Within a search universe of a given network of information tools it would be optimal to have a menu at the disposal of the users of the information tools to help these users with the identification of the data available in that search universe.

7.3 Satisfying information needs

Providing assistance to users and answers to relevant questions is at the heart of putting any information tool to use. In a search universe, users will have access to different information tools in a network and they will have different options to ensure that their information needs can be satisfied. To satisfy their information needs users will have the option to access additional support tools, drill-down to more detail, filter information or link to additional sources. Expanding their access to different tools does not only serve to satisfy the original information need but it often leads to the retrieval of additional information that can add the satisfaction of a cognitive function to the search. Because the additional information is available on demand and users are not automatically immersed into such an overwhelming information pool the user is never exposed to the danger of an overload of information.

Within such a search universe an e-dictionary becomes an information tool amongst a plethora of information tools. Varantola (2002: 38) describes this as a "modular network with seamless connections between the modules". In practice, this network allows a portal integration that allows movement from dictionary to dictionary, from dictionary to other sources and, very important, from other sources to a dictionary. In a modular network of information tools that constitutes a mutual search universe, a given dictionary is not only a source from which users move on to other sources. It can be a first stop, a transit stop, but also a non-initial and final stop for users executing search procedures to satisfy their information needs. In this regard an e-dictionary should also be a source that can be targeted by users of any other tool in the network to provide additional information on demand.

As already indicated, the demand for additional information to satisfy a specific need may lead to the satisfaction of needs not previously expressed. When obtaining information to satisfy a text reception need the extent of the accessed information may be of such a nature that it satisfies an existing albeit not previously expressed cognitive need. In this regard a second phase cognitive function is satisfied by the specific information search.

To ensure an optimal use of a network of information tools the search universe needs to be planned in such way that a user can access any tool from any other tool in the network. It should not only be mono- or even bi-directional access but rather multi-directional access.

A linear continuum presupposes a linear movement between information sources:

- From A to B to C
- From Z to Y to X

This implies that any single source can only be directly linked with its linear neighbouring sources. As one of the tools on a linear continuum a given dictionary would not allow its user an optimal access to the search universe but only to two sources. This would seriously hamper the application of the data-pulling structure of such a dictionary and would neutralise any quest for information on demand. A linear continuum is therefore not an efficient way of visualising such tools.

When visualising the network of information tools that constitutes a comprehensive search universe the information retrieval structure should preferably have a circular representation as application domain, as seen in Figure 22:

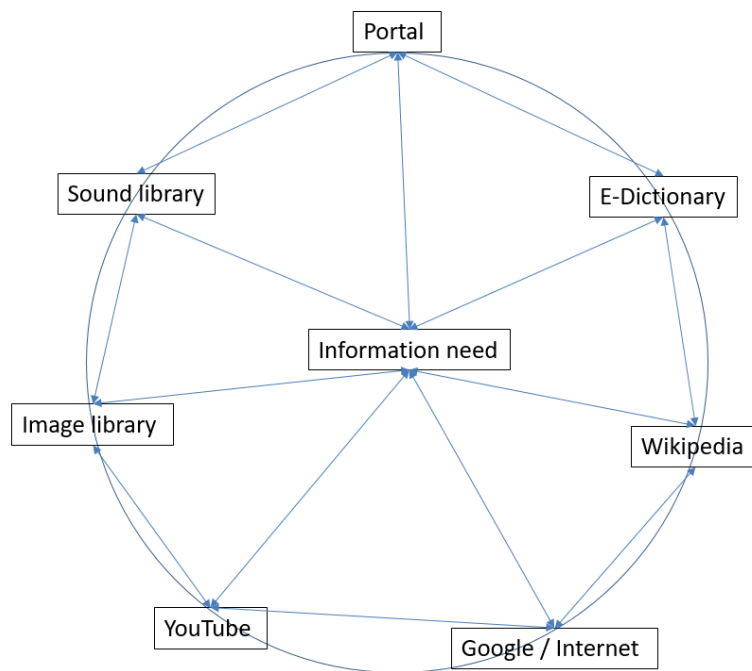


Figure 22: A modular network

Such a modular network puts the needs of the user at the centre, and the user equidistant from all information tools. From this hub the user can move along any spoke to a specific source. Users can also move between any two sources although the complexity of all possible linkages has not been indicated in Figure 22. Moving between sources can either be done directly or via the hub where the specific need is again established. The network includes a slot for e-dictionaries, some sources (like *Wikipedia*) that are not dictionaries but are "... utility tools with formal properties of a lexicographic nature" (Wiegand 2013: 285). Like dictionaries these sources could also contain curated data. This network can also include devices for user support, e.g. writing assistants. In the network one also finds sources of a non-lexicographic nature, including sources that contain non-curated data. The information need of a user will determine the source or sources they consult. Such a network is an ideal search universe for an information-on-demand approach where a user can move between sources.

When retrieving information from such a collection of information tools both pushing and pulling methods can be employed and the user should know what to use when and where. This demands not only dictionary-using skills but also a more general set of information literacy skills. These skills fall within a new kind of literacy that needs to be required, i.e. dictionary and information literacy, that will take user requirements into account.

8. In conclusion

A modular network of information tools in the e-environment has implications for lexicography. However, such a network and its different information tools should not only be discussed and analysed from a lexicographic perspective. Related fields like information science also need to contribute to a comprehensive discussion of these networks, the tools allocated to such a network, the information types populating these tools, the potential target users of one or more of these tools and the needs and information literacy skills of these users. This still needs to be done.

From a lexicographic perspective, it is clear that the dynamics of the e-environment compels metalexigraphers to carefully consider the nature and extent of the relevant lexicographic processes. It is also important to realise that lexicographers need to collaborate with experts from other fields and that the traditional notion of a dictionary culture has to be substituted by a more general information literacy culture which includes dictionary literacy.

A new assessment of users and their needs is necessary. Within the current model of dictionary research, research into dictionary use is regarded as one of the four central sections; cf. Wiegand (1998: 114). This component will have to be expanded to account for dictionary users having to use other information tools functioning along with a given dictionary in the search universe of

a network of information tools. This should also lead to the emergence of a higher level of both dictionary literacy and information literacy.

A modular network of information tools requires the user to apply their mind in all cases. This is evident from the examples discussed earlier: the user often has multiple opportunities to access or select information based on a single choice or consecutive choices they have made. An incorrect choice of meaning from the choices at hand will lead to an incorrect solution to the information need of the user, even in the case of curated information. The set of tools is not aware of the context of the information need, and only the user is aware of the context. Such a network of information tools is therefore not intelligent or "smart", but requires the user to analyse the context and make an informed choice. To be truly "smart", any set of tools should be able to provide the user with only the information that is required, as already formulated by Haas in 1962 in the context of dictionaries: "The perfect dictionary is one in which you can find the thing you are looking for preferably in the very first place you look" (Haas, 1962: 48). This level of sophisticated retrieval is currently not feasible and requires further technological developments, such as context-aware search engines and the implementation of artificial intelligence, neural networks and machine learning. Collaborative inter- and multi-disciplinary research is required to develop such systems. As Tarp and Gouws formulated this in their 2019 article: "Modern-day lexicographers are in a position to make some of the unfulfilled dreams of the past a reality. The challenge of the future is to make the impossible possible. We have work to do" (Tarp and Gouws 2019: 266). Only then will genuinely smart dictionaries and smart information tools become available.

Endnote

1. Where the term *information* is used in this paper with regard to dictionaries it covers the meaning of both *data* (= what is put into the dictionary by the lexicographer) and *information* (= what the user retrieves from a dictionary).

Acknowledgement

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant specific unique reference numbers (UID) 85434 and 95925). The grantholders acknowledge that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF supported research are that of the authors, and that the NRF accepts no liability whatsoever in this regard.

References

- Ball, Liezl H. and Theo J.D. Bothma.** 2018. Establishing Evaluation Criteria for e-Dictionaries. *Library Hi Tech* 36(1): 152-166.
- Bergenholtz, Henning and Heidi Agerbo.** 2017. Types of Lexicographical Information Needs and their Relevance for Information Science. *Journal of Information Science Theory and Practice* 5(3): 6-16.
- Bergenholtz, Henning, Theo J.D. Bothma and Rufus H. Gouws.** 2015. Phases and Steps in the Access to Data in Information Tools. *Lexikos* 25: 1-30.
- Bergenholtz, Henning, Sven Tarp and Herbert Ernst Wiegand.** 1999. Datendistributionsstrukturen, Makro- und Mikrostrukturen in neueren Fachwörterbüchern. Hoffmann, Lothar et al. (Eds.). 1999. *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft/Languages for Special Purposes. An International Handbook of Special-Language and Terminology Research, Bd./Vol. 2: 1762-1832.* Berlin: De Gruyter.
- Bothma, Theo J.D.** 2018. Lexicography and Information Science. Fuertes-Olivera, P.A. (Ed.). 2018: 197-216.
- Bothma, Theo J.D. and Henning Bergenholtz.** 2013. Information Needs Changing over Time: A Critical Discussion. *South African Journal of Libraries and Information Science* 79(1): 22-34.
- Bothma, Theo J.D., Rufus H. Gouws and Danie J. Prinsloo.** 2016. The Role of e-Lexicography in the Confirmation of Lexicography as an Independent and Multidisciplinary Field. Margalitadze, T. and G. Meladze. 2016. *Proceedings of the XVII EURALEX International Congress. Lexicography and Linguistic Diversity, Tbilisi, Georgia, 6-10 September, 2016:* 109-116. Tbilisi: Ivane Javakhishvili Tbilisi University Press. http://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202016/euralex_2016_008_p109.pdf (Last accessed on 20 April 2020).
- Bothma, Theo J.D., Danie J. Prinsloo and Ulrich Heid.** 2018. A Taxonomy of User Guidance Devices for e-Lexicography. *Lexicographica* 33: 391-422.
- Bothma, Theo J.D. and Sven Tarp.** 2014. Why Relevance Theory is Relevant for Lexicographers. *Lexicographica* 30: 350-378.
- Deolasee, Pavan, Amol Katkar, Ankur Panchbudhe, Krithi Ramamritham and Prashant Shenoy.** 2010. *Adaptive Push-Pull: Disseminating Dynamic Web Data.* <http://www-ccs.cs.umass.edu/~krithi/web/WWW10/www10/> (Last accessed on 2 April 2020).
- Duan, Zhenhai, Kartik Gopalan and Yingfei Dong.** 2005. *Push vs. Pull: Implications of Protocol Design on Controlling Unwanted Traffic.* <https://pdfs.semanticscholar.org/6f63/d37b4f8dd655e3594185e74daf4689f55aa1.pdf> (Last accessed on 2 April 2020).
- Finch, Jeremy.** 2015. *What Is Generation Z, And What Does It Want?* <http://www.fastcoexist.com/3045317/what-is-generation-z-and-what-does-it-want> (Last accessed on 28 May 2015).
- Fuertes-Olivera, Pedro A. (Ed.).** 2018. *The Routledge Handbook of Lexicography.* London/New York: Routledge.
- Fuertes-Olivera, Pedro A. and Henning Bergenholtz (Eds.).** 2011 *e-Lexicography: The Internet, Digital Initiatives and Lexicography.* London/New York: Continuum.
- Gouws, Rufus H.** 2017. La sociedad digital y los diccionarios. Domínguez Vázquez, María José and Maria Teresa Sanmarco Bande. 2017. *Lexicografía y didáctica:* 17-34. Frankfurt: Peter Lang.
- Gouws, Rufus H.** 2018. 'n Leksikografiese datatrekkingstruktuur vir aanlyn woordeboeke. *Lexikos* 28: 177-195.
- Gouws, Rufus H.** 2018a. Expanding the Data Distribution Structure. *Lexicographica* 34: 225-237.

- Gouws, Rufus H.** (to appear) Expanding the Use of Corpora in the Lexicographic Process of Online Dictionaries. Piosik, M., J. Taborek and M. Woźnicka (Eds.). 10. *Kolloquium zur Lexikographie und Wörterbuchforschung. Korpora in der Lexikographie — Stand und Perspektiven*, 19.–20. Oktober 2018, Adam-Mickiewicz-Universität in Poznań, Polen. *Lexicographica: Series Maior*. Berlin/New York: De Gruyter.
- Gouws, Rufus H.** (to appear a) 'n Verbeterde leksikografiese dataverspreiding- en inligtingsonttrekkingstruktuur. *SPIL-Plus* 2020.
- Gouws, Rufus H., Ulrich Heid, Wolfgang Schweickard and Herbert Ernst Wiegand (Eds.)**. 2013. *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin: De Gruyter.
- Granger, Sylviane and Magali Paquot (Eds.)**. 2012. *Electronic Lexicography*. Oxford: Oxford University Press.
- Haas, M.R.** 1962. What Belongs in a Bilingual Dictionary? Householder, Fred W. and Sol Saporta (Eds.). 1962. *Problems in Lexicography. Report of the Conference on Lexicography held at Indiana University, November 11–12, 1960*: 45-50. Bloomington: Indiana University.
- Hausmann, Franz J. and Herbert Ernst Wiegand**. 1989. Component Parts and Structures of General Monolingual Dictionaries. Hausmann, Franz J. et al. (Eds.). 1989–1991. *Wörterbücher. Ein internationales Handbuch zur Lexikographie/Dictionaries. An International Encyclopedia of Lexicography/Dictionnaires. Encyclopédie internationale de lexicographie*: 328-360. Berlin: Walter de Gruyter.
- Kammerer, Matthias and Herbert Ernst Wiegand**. 1998. Über die textuelle Rahmenstruktur von Printwörterbüchern. Präzisierungen und weiterführende Überlegungen. *Lexicographica* 14: 224-238.
- Müller-Spitzer, Carolin**. 2013. Textual Structures in Electronic Dictionaries. Gouws, Rufus H. et al. (Eds.). 2013: 367-381.
- Parker, Phil**. 2013. *Do You Know How Generation Z Pupils Learn?* <https://www.sec-ed.co.uk/blog/do-you-know-how-generation-z-pupils-learn/> (Last accessed on 3 April 2020).
- Tarp, Sven and Rufus H. Gouws**. 2019. Lexicographical Contextualization and Personalization: A New Perspective. *Lexikos* 29: 250-268.
- Karantola, Krista**. 2002. Use and Usability of Dictionaries: Common Sense and Context Sensibility? Corréard, Marie-Hélène (Ed.). 2002. *Lexicography and Natural Language Processing. A Festschrift in Honour of B.T.S. Atkins*: 30-44. s.l.: EURALEX.
- Wiegand, Herbert Ernst**. 1998. *Wörterbuchforschung. Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie*. Berlin/New York: De Gruyter.
- Wiegand, Herbert Ernst**. 2013. Gedruckte Gebrauchsgegenstände mit lexikographischen Formeigenschaften. *Lexicographica* 29: 285-307.
- Wiegand, Herbert Ernst, Sandra Beer and Rufus H. Gouws**. 2013. Textual Structures in Printed Dictionaries. An overview. Gouws, R.H. et al. (Eds.). 2013: 31-73.